Simplifying the use of ants as bioindicators on mine sites

Conor Robert Miller Cooper

MSc (by research)

University of York

Biology

September 2018

Abstract

Biological indicators are vital to the monitoring of ecosystems and environmental conditions across the globe as representatives of broader ecological trends. In Australia, ants are widely employed as biological indicators, owing to their ubiquity, importance to ecosystem service provision, representativeness of broader ecological patterns and well-characterised disturbance response. Ants are also considered much simpler to sample and sort than alternative indicators. However, despite these advantages, the use of ants as indicator taxa remains time-consuming, costly, and inaccessible to non-specialists due to the difficulties of identifying Australia's hyper-diverse and hyper-abundant ant fauna to species level, which limits their implementation in monitoring programmes and the research avenues that can be explored.

The drawbacks of using ants can be addressed through the use of simplified analyses which circumvent the need to use species abundance data by utilising higher taxa, restricted species lists or presence/absence data. In this thesis I analyse data from a long-term study of ant community change after mine-site rehabilitation at German Creek and Callide mines in Australia. I test four simplified analyses – Genera Abundance, Functional Group Abundance, Large-Bodied Abundance and Species Presence/Absence – in order to assess their suitability as a surrogate for species abundance data in the monitoring and evaluation of rehabilitated mine sites, by evaluating their ability to replicate key aspects of the results from a full species abundance assessment.

I found the performance of the four simplified analyses to be variable between the two mines, with the exception of Species Presence/Absence, which was able to consistently replicate key aspects of the species abundance assessment. I discuss the possible analytical and ecological factors which likely contribute to variation in performance of the four analyses and recommend a context-based approach to simplified analysis use and research, and discuss how this will enhance the use of bioindicators for monitoring environmental systems.

List of Contents

List of Figures

Chapter 2 – An evaluation of rehabilitation success at German Creek Mine & Callide Mine

Chapter 3 – Assessing the efficacy of four simplified analysis protocols for evaluating rehabilitation success

Appendix 1 – Functional Group Relative Abundance Profiles

Appendix 2 – Reciprocal Transformations of Estimated Years to Convergence

List of Tables

Chapter 2 – An evaluation of rehabilitation success at German Creek Mine & Callide Mine

Methods:

Chapter 3 – Assessing the efficacy of four simplified analysis protocols for evaluating rehabilitation success

Methods:

Results:

Chapter 4 – General Discussion

Acknowledgements

Science is a collective endeavour, and when we extend the read of human knowledge, we do so standing on the shoulders of giants. And it is in that spirit I would like to thank the giants who have made this thesis possible. First, and foremost, my supervisors; **Elva Robinson**, the model of a graduate supervisor - If I ever make something of myself in research, it will be down to what I've learned as her student - and **Ben Hoffmann**, who has been an indispensable font of knowledge for a novice myrmecologist like myself and, alongside **Alan Andersen**, & their collaborators, was responsible for the fieldwork and data collection, without which this thesis would not exist. I was also privileged to receive input from a number of other senior colleagues, who generously shared their knowledge and time with me and made important contributions to the project; **Colin Beale**, who sat on my Thesis Advisory Panel and provided many insightful comments during the project's development; **Alan Andersen**; **Chris Thomas** & the **Whole Organism Ecology Group** for their advice and feedback on this work; **Emma Rand**, for her assistance with the development of ARC-Dissimilarity analysis and for providing me with gainful and educational employment as a statistics and R-coding demonstrator, so I didn't need to get a real job; **John Pitchford**, for acting as my interim supervisor when I first arrived in York, without which I'd have been very lost; & **Bill Sutherland** and his research group at the UCCRI for their support and advice on corrections. I was also fortunate enough to start my research career in the best lab in the world, and so I must pay tribute to the Ant Lab, small but mighty, to **Dominic Burns, Eleanor Drinkwater, Josie Monaghan, Megan Holgate, Ryan Bromwell**, and to **Frances Drachenberg** & **Paula Fernández** (who were members in office if not in lab meeting) who helped me, commiserated with me, and made me feel at home for seven wonderful months; special mention must go to **Ryan Bromwell**, who understands far more about maths than I ever will, for useful discussion of statistical approaches; **Dominic Burns**, for invaluable guidance in unravelling the mysteries of R and coding Mixed Linear Models; and **Eleanor Drinkwater**, who took me under her wing and showed me the ropes of grad life with a smile. I was also lucky enough

to spend two months in the *other* best lab in the world, on the opposite side of the globe; and I'd like to thank **Ellie Kosta**, **Brian Tippett** & **Magen Pettit** at CSIRO for making me feel at home when home was so very far away, and my labmates **Juliette Chassain** & **Thomas Wind**, my companions at the very end of the Earth. I'd also like to thank **Edward Lavender** for helpful discussions about the project and always cheering me on to new heights, and **Chloe Rice** for always believing the best of me. The last year has been a wild ride, but none of it would have been possible without the MSc course, funding and support provided by the **University of York**, including my research budget, desk space and, briefly, accommodation, and by **CSIRO**, who graciously flew me out to Australia (and back), subsidised my accommodation and provided me with desk space. Last, but most definitely not least, I'd like to thank my family; **Carol Miller**, **Julian Cooper** and **Isabelle Cooper**, without whom I could never have come so far or achieved so much, who've gamely supported me galivanting off around the country and the world, and then coming back only to lock myself in my room for the rest of the summer to write. Everything I have achieved, I owe to them being at my back, always on my team, and for that I am truly grateful, for I could do none of it without their eternal love and support. To all of you, thank you.

Author's Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Results from this research project were presented in a poster at the IUSSI World Conference 2018 (Conference Poster Presentation: Surpassing our Limits: Optimising the use of ants as bioindicators on mine sites, Cooper, C., Hoffmann, B.D., Robinson, E.J.H. (2018) XVIII International Union for the Study of Social Insects International Conference, Guarujá, Brazil).

Chapter 1 – A General Introduction

1 – What are bioindicators?

The use of organisms as indicators of environmental conditions is a concept that likely has origins stretching back millennia, but it was not until after the industrial revolution that formal, scientifically rigorous indicator concepts were codified (Cairns & Pratt, 1993). Formal study of bioindicators got its start in the field of limnology, the study of inland aquatic ecosystems, with studies on either side of the Atlantic independently giving rise to the indicator species concept (Cairns & Pratt, 1993). In the USA, the concept originated in the work of S.A. Forbes on the benthic fauna of the Illinois River, beginning in the 1870s (Cairns & Pratt, 1993), and in Europe, it originated in the work of Kolwitz & Marsson (1908, 1909, cited in Cairns & Pratt, 1993) on the Saprobien system of water contamination (Cairns & Pratt, 1993). Today, bioindicators are a well-established concept applied to a wide variety of environmental assessment objectives, ranging from their original use as monitors of pollution and contamination to use in more generalised environmental assessment, monitoring of disturbance and restoration, and ecosystem management and conservation (Cairns & Pratt, 1993, Andersen & Hoffmann, 2003a, Siddig *et al*, 2016). However, much like the concept of a "species", the concept of a bioindicator, and its many synonyms, is one that may seem straightforward, but actually has a variety of definitions, each with subtle differences in meaning (Heink & Kowarik, 2010, Siddig *et al*, 2016). In their review of definitions used in the literature, Heink & Kowarik (2010) found a wide variety of definitions, ranging from very narrow to very broad, some of which were mutually incompatible. In order to resolve this, they propose the following broad, overarching definition to encompass these varied definitions:

"An indicator in ecology and environmental planning is a component or a measure of environmentally relevant phenomena used to depict or evaluate environmental conditions or changes or to set environmental goals. Environmentally relevant phenomena are pressures, states, and responses as defined by the OECD (2003)."

Under this overarching umbrella definition, Heink & Kowarik identified two major points of division in definitions:

- Measure versus Component
- Descriptive versus Normative

Indicator components are objects or processes that comprise the system, e.g. species or a fire regime, whereas measures are quantities or properties of those components, e.g. vegetation biomass, species richness or oxygen content of streams (Heink & Kowarik, 2010). Measures and components are then used in either descriptive or normative contexts. A descriptive indicator is used to describe environmental states or changes as they are, whereas a normative indicator includes a value judgement and is used in the evaluation of environmental states and changes with reference to objectives (Heink & Kowarik, 2010). Authors often do not distinguish between descriptive and normative indicators, and indeed the distinction is primarily one of purpose rather than mechanics, but it is philosophically important and provides a useful distinction between the use of bioindicators in value-neutral descriptive studies of ecology and evaluative studies for environmental planning (Heink & Kowarik, 2010). For the sake of clarity, I will adopt this broad definition of bioindicators coined by Heink & Kowarik (2010) and make use of their sub-definitions in this work, identifying the indicator component, indicator measurement and whether measurement of the indicator component is being used in a descriptive or normative context.

2 – Ants as indicator components: What makes a good indicator taxon?

Though an indicator component can refer to any component of an ecosystem, organisms are frequently selected as the component of choice. In aquatic ecosystems around the world, benthic macroinvertebrates have been a mainstay since the beginning, but until quite recently in terrestrial ecosystems plants were the predominant indicator component taxa, with fauna, particularly invertebrates, largely neglected (Andersen *et al*, 2004, Andersen & Majer, 2004, Burger, 2006). One notable exception to this trend has been in Australia, where ants were first employed as indicator components in the mid-1970s, by Majer and colleagues at the Alcoa World Alumina Australia bauxite mining operations in the Jarrah forests of southwestern Australia (Majer, 1983, Andersen & Majer, 2004, Majer *et al*, 2013). Today, ants are utilised as indicator components in a wide variety of land-management and ecosystem monitoring contexts in Australia, particularly monitoring of the impact of disturbances such as fire, grazing and mining, and the restoration of communities post-disturbance (Andersen & Hoffmann, 2003a, Andersen & Majer, 2004, Majer *et al*, 2013, Andersen *et al*, 2014, Lawes *et al*, 2017).

The popularity of ants as indicator components in Australia is a testament to their effectiveness, and ants make for particularly effective indicator components, both globally and in Australia in particular. First and foremost, ants are themselves a key faunal group in Australia, particularly in the arid zone, being uniquely diverse and abundant on this continent (Andersen, 1990). Being so abundant and diverse, ants exert great influence on a wide variety of ecosystem processes and other faunal groups (Andersen, 1990). This includes, but is not limited to, ecosystem functions such as energy and nutrient cycling and litter decomposition through scavenging, and soil formation, structuring, aeration and drainage through their nests, close associations with plants such as pollination, but particularly the harvesting and subsequent dispersal of seeds, and regulation and influencing of the populations of other faunal groups through their interactions as competitors, predators, and prey (Majer *et al*, 1982, Andersen, 1990). So ants are not only a large component of Australian ecosystem biomass

in terms of sheer abundance, but are also a critical component of their healthy functioning. This means that ants are a good choice of indicator component in Australia because they are more often than not one of the most important components of any particular ecosystem in Australia, making them worth monitoring for their own sake.

Of course, when selecting an indicator component to monitor community and ecosystem health and responses to disturbance, it is generally inherent to the selection process that the component indicates the status of other components. In this regard too, ants are a great indicator choice. In the first instance, their intrinsic importance to ecosystem function provision means that other taxa and processes are at least partially dependent on them and so the responses of these dependent taxa are likely to correlate with those of ants to some extent. In the Jarrah forests of Western Australia, species richness of ants has been found to be significantly correlated with that of total invertebrate species, and the abundances a wide variety of other invertebrate taxa, including insect larvae, Acarina, Araneae, Blattodea, Caelifera, Coleoptera, Curcilionoidea, Gryllacridoidea, Homoptera & Tettigonoidea (Majer, 1983), as well as with plant species richness, biomass of native vegetation and time since rehabilitation, with ant species diversity correlated with total vegetation biomass (Majer *et al*, 1982). Ant diversity was also positively associated with litter cover (Majer, Brennan & Moir, 2007). Their post-disturbance responses have not been found to correlate so closely with vertebrate taxa, however (Fox, 1982, Nichols & Nichols, 2003). That said, the community composition of ants has been found to be strongly associated with overall ecosystem community composition, even when vertebrates are included (Bisevac & Majer, 2002). Overall, trends in ant species richness and diversity show unusually strong ties to a wide variety of invertebrate taxa and to plant diversity and biomass, as well as broad community composition, making them a strong candidate for selection as an indicator component in terms of representativeness (Majer, Brennan & Moir, 2007).

Representativeness is important for a potential indicator component taxon, but equally important is a well-characterised disturbance response. Here ants once again come out on top, having one of the best-characterised responses to disturbance of any invertebrate group (Andersen & Hoffmann, 2003a). This characterisation covers responses to mine rehabilitation, fire, grazing, pollutants, agriculture, urbanisation, and even the impact of military exercises (Woinarski *et al*, 2002, Andersen & Hoffmann, 2003a). The response of ants to post-mining rehabilitation is particularly well-studied across Australia and beyond (Fox, 1982, Majer *et al* 1982, Majer, 1983, Majer, 1997, Jackson & Fox, 1996, Bisevac & Majer, 1999, Andersen, 1997, Andersen *et al*, 2002, Andersen, Gómez *et al*, 2003, Hoffmann & Somes, 2003, Nichols & Nichols, 2003, Andersen & Hoffmann, 2003a, van Hamburg *et al*, 2004, Ottonetti, Tucci & Santini, 2006, Ribas *et al*, 2012).

Finally, ants also hold practical advantages over other invertebrates as an indicator taxon. This is due to the fact that they are easily and simply sampled, using simple equipment such as pitfall traps, as most Australian ants are epigaeic foragers and so the use of pitfall trapping targets most of the community (Andersen, 1990, Greenslade, 1979, cited in Andersen, 1990, Alonso & Agosti, 2000) Ants are also comparatively easily sorted once sampled compared to other invertebrate taxa. At the species level, as a rule all samples are adult female workers so identification is not complicated by sexual or age-based polymorphisms, and, as endopterygotes, ants lack immature instars which could confound sorting, unlike other candidate indicator species such as soil mites (Andersen, 1990, Cuccovia & Kinnear, 1997, Nakamura, 2003). Identification is also supported by the taxonomy being based on external morphological features rather than reproductive organs (Andersen, 1990). The task of sorting is even simpler at the level of genera, with differences between genera being so distinctive that classification can be done with minimal effort and even in the field, in contrast to other invertebrate groups, where differentiation between genera requires specialist knowledge (Andersen, 1990). These features make ants much easier to work with than other invertebrates, with much less time,

effort and specialist knowledge needed to use them as indicator components than other taxa, and so a much more feasible choice for integration into monitoring programmes than potential alternatives (Andersen, 1990).

Ants therefore are a key taxon in Australian ecosystems which is also representative of a much broader range of taxa, well-studied and characterised in terms of disturbance response and relatively easy to work with. These qualifications have secured the place of ants in terrestrial ecosystem monitoring in Australia. However, while ants are the best choice of indicator taxon, they are still not an easy choice of indicator taxon. Sorting of samples remains an arduous, time-consuming task, and although less complicated than the identification of other invertebrates, still frequently requires specialist knowledge to identify each of the hundreds or even thousands of samples to species. Furthermore, the taxonomy of Australian ants is very incomplete, so samples may not be identifiable to species and must instead be assigned a study-specific species code (Andersen & Hoffmann, 1998a). These factors restrict the use of ants in monitoring and management efforts, as their use still requires considerable time, expense and specialist knowledge, so finding ways to address these issues would yield great social, economic and environmental benefits.

3 – Mine site rehabilitation in Australia

Mining has been part of Australia's economy for 200 years and is one of Australia's most important industries, comprising 9% of country's total GDP and 50% of its exports (Minerals Council of Australia, 2017). However, the cost of this productivity is an incredibly environmentally destructive extraction process.

Mining companies are currently committed by federal and state government legal and internal industry regulatory frameworks to rehabilitation of mine sites at the conclusion of mining activities (Minerals Council of Australia, 2017). Over time, rehabilitation practices and requirements have evolved. The long-studied site of Alcoa of Australia Ltd. Western Australian bauxite mines highlights this process of evolution. The requirements for new mines at the Alcoa site have

evolved from a commitment to leaving the area tidy under the first agreement in 1961, to a commitment to some rehabilitation in the form of reforestation and prevention of soil erosion in 1972, and by 2012 Alcoa had been committed to continuous research, monitoring and evaluation of its rehabilitation practices, with the aim to restore exhausted mining pits to the original Jarrah ecosystem featured at the site (Majer *et al*, 2013). The rehabilitation procedures employed have evolved in step, progressing from planting non-native species in 1966 to gradual improvements in practice throughout the 1970s, including ripping of the mine floor to improve root penetration, the use of fresh rather than stockpiled topsoil, direct seeding of native understorey species and two-layered removal of topsoil to preserve the seed bank and nutrient layers of the topsoil (Majer *et al*, 2013). A similar stepwise improvement of rehabilitation methods occurred at the Allied Eneabba Ltd and Associated Minerals Consolidated Ltd. Eneabba Mining area. In the 1970s, Majer reported that experimental small-scale rehabilitation at the Eneabba mining area, Western Australia, took the form of re-spreading of topsoil and associated plant material collected prior to mining over mining tailings after the completion of mining up to two years later, and later the application of mulched vegetation from the surrounding heathland over the topsoil (Majer *et al*, 1982). By 1998, a much more sophisticated programme had been implemented, with topsoil now sourced from new mining developments where possible and removed in such a way as to preserve the soil seed bank, ensuring the seeds and mycorrhiza of the topsoil being added were viable, and an extensive seed collection and propagation programme having been established to re-establish those species that cannot be restored from the seed banks of the topsoil or mulch (Bisevac & Majer, 1999).

Part of the monitoring and evaluation of the restoration process post-rehabilitation at these sites and others has been the inclusion of ants as indicator taxon, as discussed in section 2 of this introduction. Studies of restoration of mine sites using ants as indicator components typically make use of a comparative or convergence-based approach, where measurements of the ant community such as species richness and diversity, or community composition,

from rehabilitated sites are compared with representative local undisturbed sites (Majer *et al*, 1982, Andersen *et al*, 2003, Hamburg *et al*, 2004). Although some studies monitor the same sites over an extended period of time (Andersen, Hoffmann & Somes, 2003, Majer & Nichols, 1998, Majer *et al*, 2013), most studies have been single-instance studies surveying sites of a range of ages at once and constructing a chronosequence (see Majer & Nichols, 1998) from the results (Andersen & Majer, 2004). The cost limitations of surveying ant species abundance (as discussed in section 2) have likely contributed to the lack of long-term studies. Despite the fact that the Alcoa bauxite mines were some of the first rehabilitated sites to have been studied using ants, and have continued to be monitored, on and off, for over 37 years (Majer *et al*, 2013), the sites have still not fully recovered to a native assemblage, although there has been a noticeable improvement in how close rehabilitated sites have come to restoration as rehabilitation methods have improved (Majer *et al*, 2013). To the author's knowledge, no Australian rehabilitated site studies to date have discovered a rehabilitated site with an ant community composition fully resembling that of nearby native vegetation, with the exception of at Callide Mine, where a single site (TH91) was discovered to have a community composition closely resembling that of the reference sites (Andersen, Hoffmann & Somes, 2003). This demonstrates both the long time-lag of recovery after mining and the importance of further research and study into ecosystem rehabilitation, as almost 50 years of mine-site rehabilitation monitoring with ants has yet to yield a single definitively converged ant community. This research could in turn be facilitated by the addressing of limitations of ants as an indicator component taxon, enabling more and longer studies of post-rehabilitation ecosystem recovery.

Chapter 2 – An evaluation of rehabilitation success at German Creek Mine & Callide Mine

Introduction

Ants are widely utilised as indicator taxa in Australia, particularly in the mining industry, which has embraced them as a means of monitoring and evaluating their rehabilitation efforts (Andersen & Majer, 2004). The systems developed in Australia are also increasingly being embraced worldwide, with the mining industry once again leading the way, as pressure grows for ecologically sustainable development, and scientists and agencies outside Australia recognise the value of incorporating terrestrial invertebrates into monitoring efforts (Majer, 1997, Gómez *et al*, 2003, Andersen & Majer, 2004, van Hamburg *et al*, 2004, Ottonetti, Tucci & Santini, 2006, Ribas *et al*, 2012, Grandin *et al*, 2013, Siddig *et al*, 2016).

However, the majority of studies utilising ants as indicator taxa remain short-term affairs that utilise a "chronosequence" approach (*sensu* Majer & Nichols, 1998), where a range of sites of a range of ages are surveyed and an artificial time sequence of ant succession is constructed from those surveys, as a substitute for long-term monitoring of individual sites (Majer & Nichols, 1998, Andersen & Majer, 2004). The use of chronosequence approaches instead of long-term monitoring represents a trade-off. Chronosequences offer the capacity for replication and hence quantification of factors affecting restoration success, whereas long-term studies are comparatively costly and so are restricted to low or no replication, meaning the results are not necessarily applicable to other sites (Majer & Nichols, 1998). However, the replication of chronosequences is limited by variation in site conditions and in rehabilitation approach – the rapid rate of evolution of rehabilitation practices means that the oldest sites often have undergone a very different rehabilitation regime than those towards the more recent end of the continuum (Majer & Nichols, 1998). These limitations

and gaps in site ages mean that it is not always possible to fully describe the ant succession with chronosequences, while long-term studies provide a continuous record of changes in the ant community from rehabilitation onwards, under fixed starting conditions (Majer & Nichols, 1998). However, in a comparison of the two approaches, the results of the chronosequence study were confirmed by the long-term results, suggesting chronosequences may be an acceptable substitute with the advantage of being much faster to perform (Majer & Nichols, 1998). However, as long-term studies are rare, and comparison between a chronosequence and long-term study of the same site has only occurred once (Majer & Nichols, 1998), it is unclear how applicable this result is to other study sites. From a management perspective, chronosequences present additional problems, as they are only an estimation of rehabilitated site community changes, and given the variable conditions, do not guarantee that newer rehabilitated sites will follow the same trajectory. Indeed, different rehabilitated sites can in fact take some very different paths following rehabilitation (see results). The somewhat abstract statistical nature of chronosequence results also made them harder to convey to non-scientists than the relatively clear trends of long-term study sites (Majer & Nichols, 1998). As such, while chronosequences appear to be an adequate substitute for long-term studies in investigations of environmental factors influencing post-rehabilitation ant community recovery, this needs further verification, and where possible long-term studies remain a more complete and comprehensible account of post-rehabilitation community dynamics, and irreplaceable for monitoring programmes.

The current study evaluates the success of the rehabilitation programmes at two mines, Callide and German Creek, in Queensland, using the ant community as a normative indicator component and community composition (abundances of species) as the indicator measurement. The study is somewhat unique in that it is a relatively short-term, but multi-year (4-6 years), study of multiple rehabilitated sites of a range of ages, combining aspects of the long-term and chronosequencing approaches. This provides uniquely precise data on ant community changes at each site in the years following site rehabilitation,

allowing for the observation of year-on-year changes in community composition during the restoration process. It is also unique in that it makes use of two sets of monitoring data, meaning that, although the data from the two mines cannot be treated as replicates or pooled due to differences in location, ecology, length of study and study time, it is possible to qualitatively assess the broader applicability of trends in the recovery process through the degree of agreement in results between mine sites. By establishing the trends in ant community composition at the level of species abundance, this study will serve as a baseline for the evaluation of various simplified analyses (see Chapter 3) for their fidelity to the trends observed at the level of species abundance.

Methods

1 – Study Sites

Data had been collected from German Creek Mine, Queensland, from 1997 to 2001, and from Callide Mine, Queensland, from 2001-2006 by Andersen & Hoffmann (except for Callide Mine in 2003, when data was collected by Stacey). (Andersen & Hoffmann, 1998a, 1998b, 1999, 2000a, 2001a, 2001b, 2002, 2003b 2005, 2006, Stacey, 2003, Hoffmann & Andersen, 2004). While the sampling and analysis procedures for both sites are the same unless specified, note that each site was subject to an independent study with independent species codes, and the data from each mine site are analysed separately.

German Creek is a black coal mine in the Bowen Basin, 25km SW of Middlemount and 240km from Mackay, in the semiarid tropics of Queensland. Land rehabilitation is a continuous process that follows half a panel behind the coal-mining operation, so the ages of different rehabilitated sites vary (Anglo American 2018a). Ants were sampled from three reference sites and six rehabilitated sites (table 1). Reference sites 1, 3 & 7 were selected as a representative subset of the local natural habitats from a 1997 pilot survey of local undisturbed habitats (Andersen & Hoffman, 1997). Reference site 1 consisted of Blue Gum Woodland, 3 of Lancewood low woodland, and 7 of Yapunyah Woodland (Andersen & Hoffmann, 1998). The six rehabilitated sites were pit dump rehabilitations ranging from 1-11 years old as of 1998. Rehabilitated sites 8 & 9 were initially sampled in the 1997 pilot survey, whereas 10, 11, 12 and 13 were sampled for the first time in 1998.

Table 1: *Summary description of rehabilitated sites at German Creek Mine (1998) (taken from Andersen & Hoffmann, 1998b)*

	Site 8	Site 9	Site 10	Sit 11	Site 12	Site 13
Date of rehabilitation	1986	1994	1993	1993	1996	1994
Age of sampling (yrs)	11	3	4	4	1	3
Soil type	Grey-brown sandy clay loam	Grey-brown clay loam	-	Red-brown sandy loam	-	Grey-brown sandy clay loam
Slope	20%	20%	10%	20%	20%	20%
Contour interval (m)	nil	50	nil	50	50	50
Dominant grass species	Golden Beard (*Chrysopogon fallax*)/ Red Natal (*Melinis repens*)	Indian Blue (*Bothriochloa pertusa*)	Buffel (*Cenchrus ciliaris*)	Buffel (*C. ciliaris*)	Red Natal (*M. repens*)	Buffel (*C. ciliaris*)

Callide Mine is a black coal mine in the Callide Basin, 20 kilometres from Biloela and 120km SW of Gladstone, in the semiarid tropics of Queensland (Anglo-American, 2018b). Land rehabilitation is concurrent to mining, taking place when a pit is mined out (Andersen, Hoffmann & Somes, 2003, Anglo-American, 2018b). Standard rehabilitation procedure involves the removal of vegetation and stockpiling of native topsoil prior to mining, and following mining overburden dumps are reshaped into stable landforms and the stockpiled soil is respread and deep-ripped before seeds of locally collected plant species are sown (Andersen, Hoffmann & Somes, 2003). Rehabilitation techniques develop over time however, and prior to 1997 sites were not deep-ripped and pasture grasses were the dominant plants sown, although these tend not to persist once trees become established (Andersen, Hoffmann & Somes, 2003). Rehabilitated site DSC81 was rehabilitated in 1981, preceding the use of topsoil (Andersen & Hoffmann, 2001). Ants were sampled from three reference sites and eight rehabilitated sites (see Table 2). Reference sites 6, 8 & 9 were selected as a representative subset of the local natural habitats from a 2000 pilot survey (Andersen & Hoffmann, 2001b). Reference site 6 consisted of Ironbark woodland with grassy understorey on rocky soil, 8 of Lancewood woodland on gravelly soil, & 9 of Gum-topped Box woodland on gravelly soil (Andersen & Hoffmann, 2001b). The eight rehabilitated sites were all spoil dumps – BHS94, BH99 and TH91 were all spoil dumps placed over natural surfaces, DCB98 was backfilled over a dragline strip, and DSC81, TGC92, DCB94 & TGB98 were dragline spoil slopes dumped on adjacent natural surfaces (Andersen, Hoffmann & Somes, 2003). The sites chosen represented a mix of ages (2-20 years old as of 2001) and rehabilitation techniques (Andersen & Hoffmann, 2001b). Sites BHS94 and DSC81 were initially sampled in the 2000 pilot survey, while all other rehabilitated sites were incorporated in 2001 (Andersen & Hoffmann, 2001b).

Table 2: *Summary description of rehabilitated sites at Callide Mine (2001) (Taken from Andersen & Hoffmann, 2001b)*

	DCB94	DCB98	DSC81	TGC92	TGB98	TH91	BH99	BHS94
Date of rehabilitation	1994	1998	1981	1992	1998	1991	1999	1994
Age (years) at sampling	7	3	20	9	3	10	2	7
Slope	15%	8%	25%	17%	16%	10%	12%	20%
Litter Depth	0cm	2cm	5cm	2cm	10cm	10-15cm	5cm	5cm
Ground Cover (%, composition)	20%, rocks, a few twigs, logs	30%, mainly grass and rocks, little amount of leaves. Limited deep-ripping	80%, mostly grass and a few leaves, scattered large rocks	30%, mainly leaves with some twigs, grass & logs	85%, grass	85%, mostly leaves, bark, twigs and grass	70%, mainly grass	85%, mostly grass with some leaves & logs

Vegetation	Acacias germinated, poor grass germination	Acacias present but restricted in sampling area. Fairly sparse cover of Red Natal (*M. repens*) with some Rhodes grass (*Chloris gayana*) & Buffel (*C. ciliaris*)	-	Patchy, primarily Buffel (*C. ciliaris*) and Rhodes grass (*C. gayana*). Some lemon-scented gum (*Corymbia citriodora*), ironbark eucalypts, acacias	High levels of introduced pasture grasses and some acacias	High vegetative cover, extremely good *Acacia* and eucalypt establish-ment and cover, and some shrub and sparse native grass understorey	Low cover of native and introduced grasses & legumes, good numbers of *Acacia* and eucalypts also present – cover not fully developed to high levels but good development conditions	Early-stage cover of green panic (*Panicum maximum*), Buffel (*C. ciliaris*), Red Natal (*M. repens*), *Urochloa* & *Acacia*, with some eucalypts becoming evident

2 – Sampling

Ants were sampled using 6.5cm diameter pitfall traps, partly filled with ethylene glycol as a preservative. At each sampling site a 5x3 grid of 15 pitfall traps with 10m spacing was established at each site (except BHS94, where only 14 traps were operative) during the wet season (Jan-Feb). At German Creek traps were operated for 4 days, whereas at Callide traps were operated for 5 days. Sampling was carried out from 1997 to 2001 at German Creek and from 2000 to 2006 at Callide. At Callide, not all sites were sampled in all years due to sampling constraints, and some further samples were excluded during analysis. For a full listing of exclusions see table 3.

Table 3: *List of samples excluded from Callide Mine analysis*

Year	Site(s)	Justification
2001	DSC81	Collected from wrong site (Andersen & Hoffmann, 2002)
2004	Ref 9, DSC81	Unusually low species richness in all 2004 samples, particularly these two, possibly affected by rain during sampling period & reliability questionable (Hoffmann & Andersen, 2004)
2005	All reference and rehabilitated sites except TH91	Decision made in this year to focus only on TH91 (Andersen & Hoffmann, 2005)

3 – Analysis

Ants had been sorted to species level and their abundances in each trap pooled to give site-level abundance for each year by Andersen & Hoffmann (Andersen & Hoffmann, 1997-2006, Stacey, 2003). The taxonomy of northern Australian ants is poorly known and most species are undescribed. As such, where possible, unidentified species had been assigned to species groups following Andersen (2000, cited in Andersen, Hoffmann & Somes, 2003) and code numbers had been assigned to each species that only apply to those studies (Andersen, Hoffmann & Somes, 2003).

Analysis was conducted by Conor Cooper in R v.3.4.3 with packages *vegan* and *nlme*.

Ant abundances were square-root transformed to increase the contribution of rare species to site dissimilarity, and particularly to down-weight the contributions of species of *Iridomyrmex* (*rufoniger* gp), which includes a number of hyperabundant species (*Iridomyrmex* species P & C (*rufoniger* gp) at Callide Mine, and *Iridomyrmex* species B & E, (*rufoniger* gp) at German Creek) with abundances orders of magnitude greater than other species at several samples, while retaining community structuring information in analysis (Clarke & Warwick, 2001).

Assessment of the recovery of rehabilitated mine-sites was carried out through the normative use of ants as an indicator component, with ant community composition (abundances of species) as the indicator measure. Recovery was assessed by comparison of ant community composition at rehabilitated sites to ant community composition at the unmined reference sites, sampled at the mines at the same time, representing the natural ant communities that existed on rehabilitated sites prior to mining. Sites with communities which become more similar to those on one or more of the reference sites are considered to be recovering, while those without are not. Though reference sites undergo fluctuation in community composition, it is assumed that they are not undergoing significant directional change in community composition, and so are suitable as a point of comparison.

In order to overcome the "moving target" nature of reference site samples, for each reference site the samples were used to estimate the Average Reference Community (ARC). These were estimated by calculating the median square-root transformed abundances of each species occurring at the site during the sampling period. Rehabilitated sites were then compared to these ARCs, rather than to the position of the reference sites in any given year. This enabled us to

calculate the extent of a rehabilitated site's convergence with the reference sites, and hence its recovery, based on its dissimilarity to the ARC (ARC-Dissimilarity approach).Dissimilarity between site community samples, including the calculated ARCs, were calculated with a Bray-Curtis dissimilarity index, and the resulting distance matrix analysed with nonmetric Multidimensional Scaling (NMDS) fitted to principle components (Ottonetti, Tucci & Santini, 2006, Clarke & Warwick, 2001).

In order to assess whether sites were converging with age, a linear model of the effect of site age on dissimilarity of rehabilitated site samples from ARCs, with Site as a random factor, was fitted to ARC-Dissimilarity outputs (BC dissimilarity from ARC scores). The tracking of changes in an index of similarity between rehabilitated and reference site ant communities over time in order to chart recovery progress has been utilised in previous long-term monitoring of rehabilitated site ant community recovery (Majer & Nichols, 1998).

An idealised recovery scenario under the convergence model would see Bray-Curtis dissimilarities of rehabilitated sites from one or more ARCs decreasing year-on-year, undergoing clear directional movement towards the ARCs in NMDS Ordinations and a significant decrease in BC Dissimilarity with age on ARC-Dissimilarity plots. Eventually rehabilitated sites would enter into a stable, non-directional pattern of fluctuation in community composition centred upon an average community composition resembling that of one of the reference site ARCs. A more generalised recovery scenario may see a rehabilitated site entering into such a pattern around an ARC position that does not resemble any one reference site but instead falls within the area of ordination space occupied by the reference sites as opposed to pre-recovery rehabilitated sites.

In order to gain insight into the changes in species composition driving convergence, changes in relative abundance of functional groups at each site were also examined. The Functional Group scheme is a classification scheme for Australian ants based on competitive interactions, habitat requirements and

responses to temperature, stress and disturbance (Andersen, 1990). This classification divides Australian ants into nine functional groups based on their competitive interactions with each other, biogeography, and tolerance to temperature, stress and disturbance: Dominant Dolichoderinae, Subordinate Camponotini, the three Climate Specialist groups: Tropical, Cold and Hot Climate Specialists, Generalised Myrmicinae, Opportunists, Cryptic Species and Specialised Predators (Andersen, 1990). Although initially used to classify ant communities on biogeographic scales, the functional group scheme has since been found to be useful for monitoring disturbance, particularly major disturbances like mining which completely transform a habitat, analogous to moving across a biogeographic boundary, with comparable effects on the functional group profile (Andersen, 1993, Bisevac & Majer, 1999, Andersen & Hoffmann, 2003). Therefore the changes in the relative abundances of each functional group in the ant species community reveal details about what changes are occurring in the makeup of the ant community and the ecological factors driving those changes (Andersen, 1995, Andersen & Hoffmann, 2003).

Results

1 – Callide Mine

1.1 – Overview

Reference and rehabilitated sites are not clearly separated in Non-Metric Multidimensional Scaling (NMDS) ordination (fig. 1), although the reference samples clearly group in the bottom-left corner of the ordination, and separation between reference and rehabilitated sites occurs predominantly on axis 1. Examination of the samples in age order shows that reference community samples do not appear to exhibit systematic directional movement in ordination with time, and so there is no evidence that they are undergoing systematic variation (Andersen & Hoffmann, 2003b). This means the key assumption of the Reference-Comparison rehabilitation assessment, that variation in reference community composition is random rather than systematic, is upheld. Of the three reference sites, 8 and 9 overlap substantially on both NMDS axes, while reference 6 overlaps with them on NMDS axis 2 but is much more centrally positioned on axis 1. As a result, the rehabilitated sites are overall much closer to reference site 6 than to the other two reference sites. This also means the convergence of rehabilitated sites to reference site 6 is much more strongly affected by movement along axis 2 than convergence on reference sites 8 or 9, and this is reflected in the ARC-Dissimilarity results (fig. 2).

Four rehabilitated sites overlap with reference site 6 on axis 1 (BHS94, DCB94, DSC81 & TGC92) at various times in the sampling period. Of these four, only BHS954 and DSC81 simultaneously have a close association on axis 2 as well, both in the last year of sampling. Only one site overlaps with references 8 & 9 on axis 1, TH91.

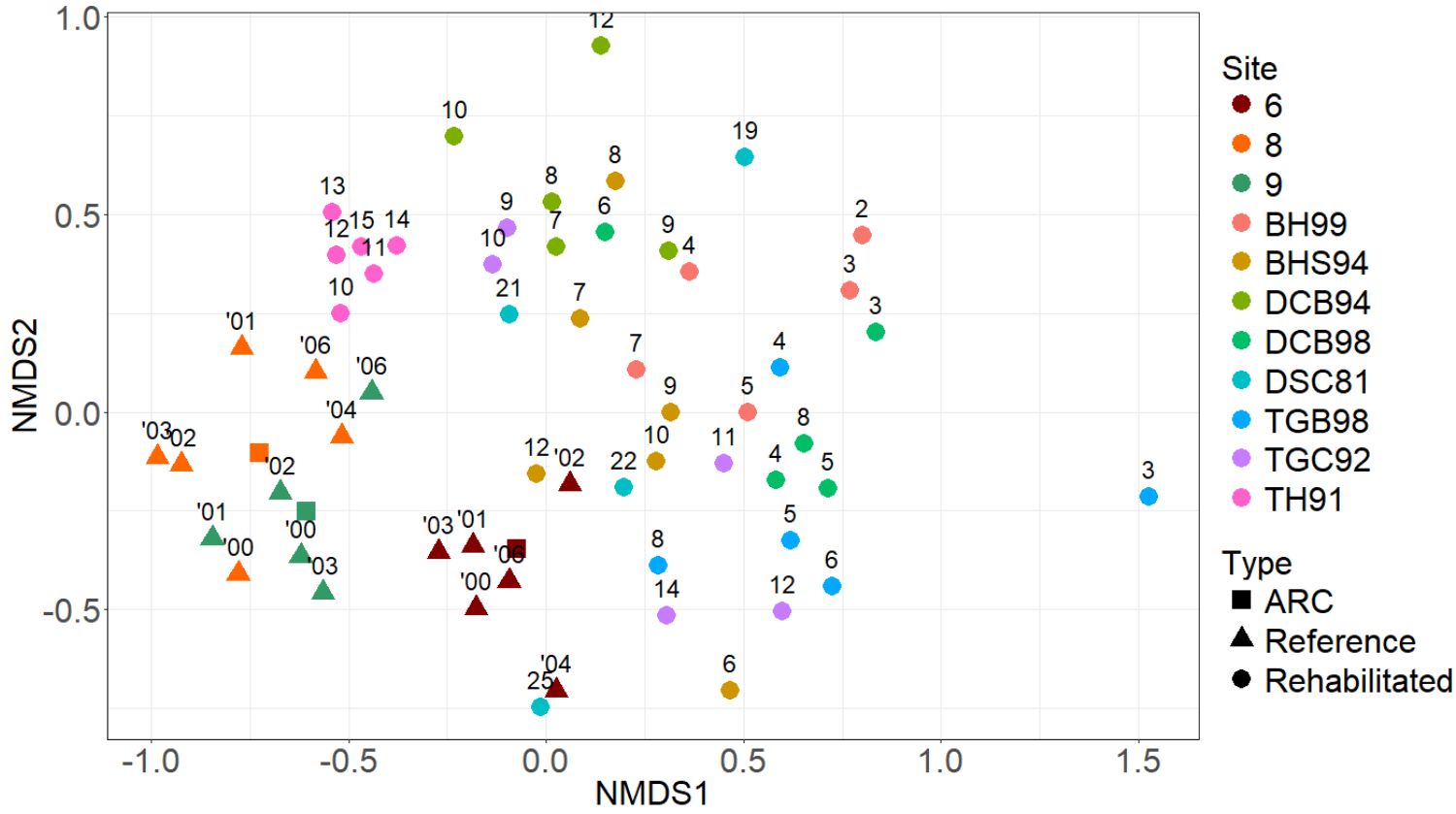


Figure 1: *Non-metric multidimensional scaling ordination of Callide Mine Sites based on ant species abundances, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (2000-2006) for reference site samples. Two-dimensional stress = 0.2291872*

Linear models of overall trends of convergence across rehabilitated sites with time showed that, as a group, rehabilitated sites significantly converged on reference sites 6 and 9, but not on reference site 8 (Site 6: slope = -0.0054, SE = 0.002, t(32) =-2.83, p = 0.008) (Site 8: slope = -0.0043, SE = 0.003, t(32) = -1.59, p = 0.121) (Site 9: slope = -0.0087, SE = 0.003, t(32) = -3.31, p = 0.002). These results suggest that overall the rehabilitation programme at Callide Mine is succeeding.

1.2 – Rehabilitated Sites

BH99 and **TGB98** show the most linear convergence path along NMDS axis 1 (fig. 1) with substantial movement along axis 1 and relatively little movement on axis 2, and this is reflected in ARC-Dissimilarity (fig. 2), where both sites are shown to be converging with all three reference sites. BH99 has a notable downward trajectory on axis 2 as well however, and so is converging substantially more rapidly on reference sites 6 and 9 than reference site 8, which occurs higher up axis 2 and so is further out from BH99's convergence trajectory.

BSH94, the third site which converges with all three references in ARC-Dissimilarity, takes a more erratic, spiralling path in the NMDS ordination. The majority of its movement is up and then down axis 2 towards reference site 6, with relatively little movement along axis 1 which is enough to push it into very close association with reference site 6 with a correspondingly rapid convergence trajectory in ARC-Dissimilarity, but yields only a relatively gradual convergence trajectory with reference site 8 and 9, as the movement towards either site by BHS94 is relatively small compared to its movement towards reference site 6.

DSC81 is the fourth site to converge on all three references in ARC-Dissimilarity, although is converging only very gradually with reference site 8. DSC81 shows substantial convergent movement in NMDS too, largely in years 1-3, while after that systematic movement is mostly on axis 2. If sampling were to have been extended this trend may have continued, DSC81 rapidly diverging from all three sites in the same manner as DCB94 (see below), but as the data currently stand,

at the conclusion of sampling this trajectory sees DSC81 rapidly converging on reference site 6 (figs. 1, 2a). This downward trajectory places it on a clear convergent path with reference 9 (2c) too, although notably DSC81 does passes by reference site 9 in NMDS, as it may potentially do for reference site 6. In contrast, this trajectory results in a very shallow decline in dissimilarity from reference site 8 (2b) overall, since ARC of reference site 8 is positioned higher up axis 2 than that or reference sites 6 or 9 and so DSC81's downward trajectory does not contribute as much to convergence with reference site 8 as it does to convergence with the other two reference sites.

DCB94 diverges from all three references in ARC-Dissimilarity (fig. 2), although is close to neutral relative to reference 8. This reflects its relative lack of overall movement on NMDS axis 1 (fig. 1), with its position at the beginning of the sampling period being close to its final position, and its systematic divergence from all three reference sites on axis 2, moving up and away from all three.

DCB98 does not move systematically in NMDS ordination (fig. 1) and appears prone to large fluctuations in community composition. This lack of systematic movement would appear to be reflected in the ARC-Dissimilarity graphs for reference sites 8 & 9 (figs. 2b & 2c), in relation to which DCB98 displays a very shallow convergent trajectory and an even more shallow divergent one respectively, but is clearly converging with reference site 6. The NMDS and ARC-Dissimilarity results do not appear to agree for this site. It is important to note that NMDS is a low-dimensional representation of a much more complex arrangement of sites, so higher-dimensional representations may correspond more closely to the ARC-Dissimilarity results.

TGC92, in ARC-Dissimilarity (fig.2), is rapidly converging with reference site 6 (2a) but is diverging from reference sites 8 & 9 (2b & 2c). This is reflected in its s-shaped NMDS trajectory (fig. 1). Since TGC92 begins in alignment with reference site 6 on axis 1, and reference site 6 is lower down axis 2 than either 8 or 9, this

trajectory causes it to diverge from all three reference sites on axis 1, but brings it down towards reference site 6 on axis 2, closer than it started. Like DSC81, the trajectory of TGC92 may go on to carry it on past reference site 6 into divergence, but as of surveying it remains convergent on this reference site.

TH91 is the only rehabilitated site to overlap with reference sites 8 & 9 on NMDS axis 1 (fig. 1), being positioned further left on the predominant axis of separation of reference and rehabilitated sites than reference site 6, and this is reflected in their ARC-Dissimilarity results (figs. 2b & 2c) which place TH91 as the least dissimilar sites to the reference in each case. It experiences the least change in position year on year, even less than the three reference sites, giving the appearance of a stable, successfully rehabilitated site. However, in ARC-Dissimilarity (fig.2) it is divergent from reference sites 6 & 8 (2a & 2b) and displays an odd "arch" shaped trend relative to reference site 9. This reflects the fact that TH91 in NMDS ordination actually shows a directional trend to its year-on-year movement, moving up and to the right, diverging from the reference sites, although it doubles back in the final year of sampling.

Overall there have been very mixed results from the Callide Mine rehabilitation programme – while some rehabilitated sites are clearly converging on reference sites and on the road to recovery (BH99, TGB98, BSH94), the status of others is more ambiguous (DSC81, TGC92, DCB98, TH91), and DCB94 appear to be actively diverging from all reference communities. Out of the four ambiguous rehabilitated sites, two may potentially be heading for divergence in the future (DSC81 and TGC92) based on their trajectories in the NMDS ordination, and two converge on a single reference site (TGC92 on site 6 and TH91 on site 9) while diverging from the other two.

Figure 2 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 6 (2a), 8 (2b) and 9 (2c) with site age since rehabilitation (years) at Callide Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 2a: *Convergence with Site 6, Callide Mine*

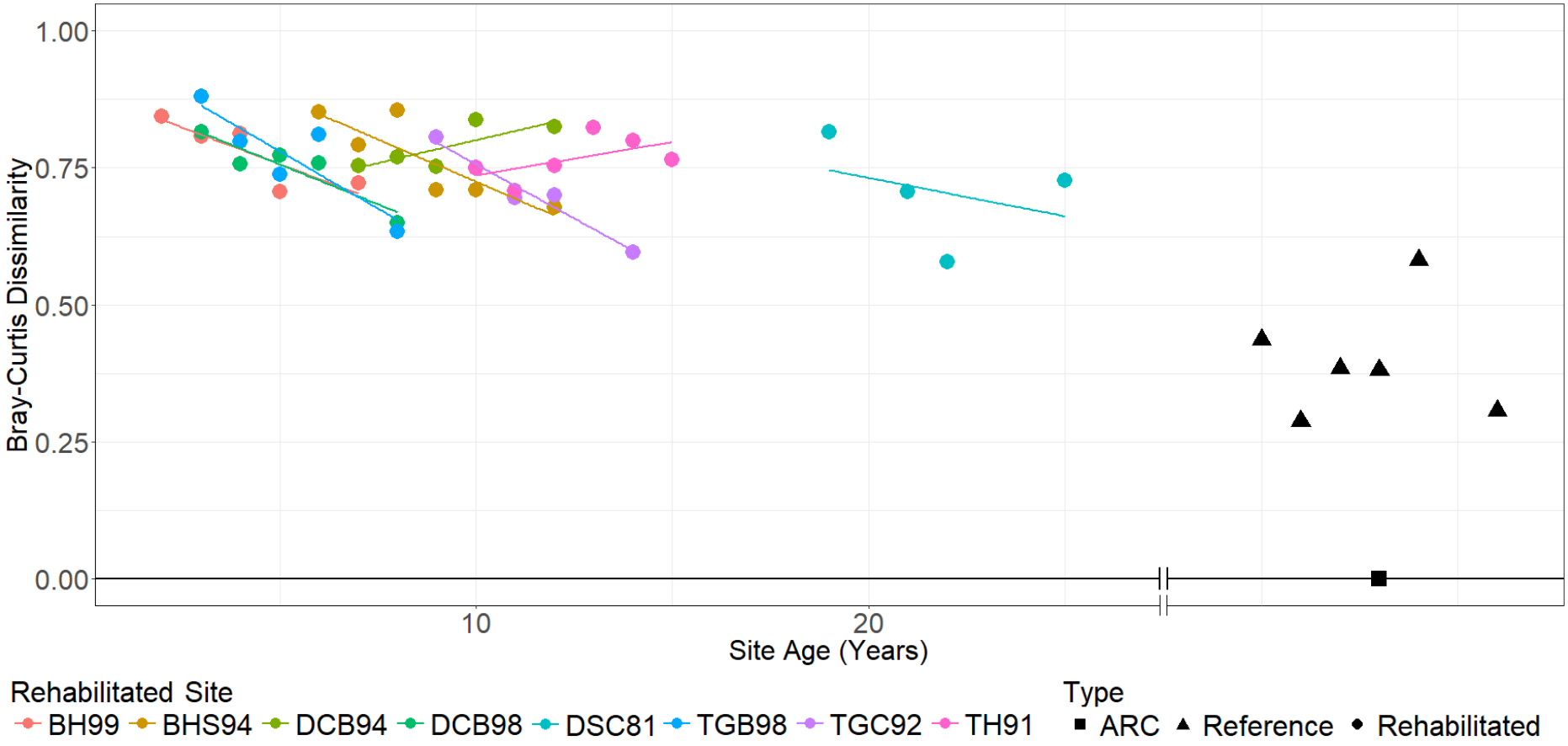


Fig 2b: *Convergence with Site 8, Callide Mine*

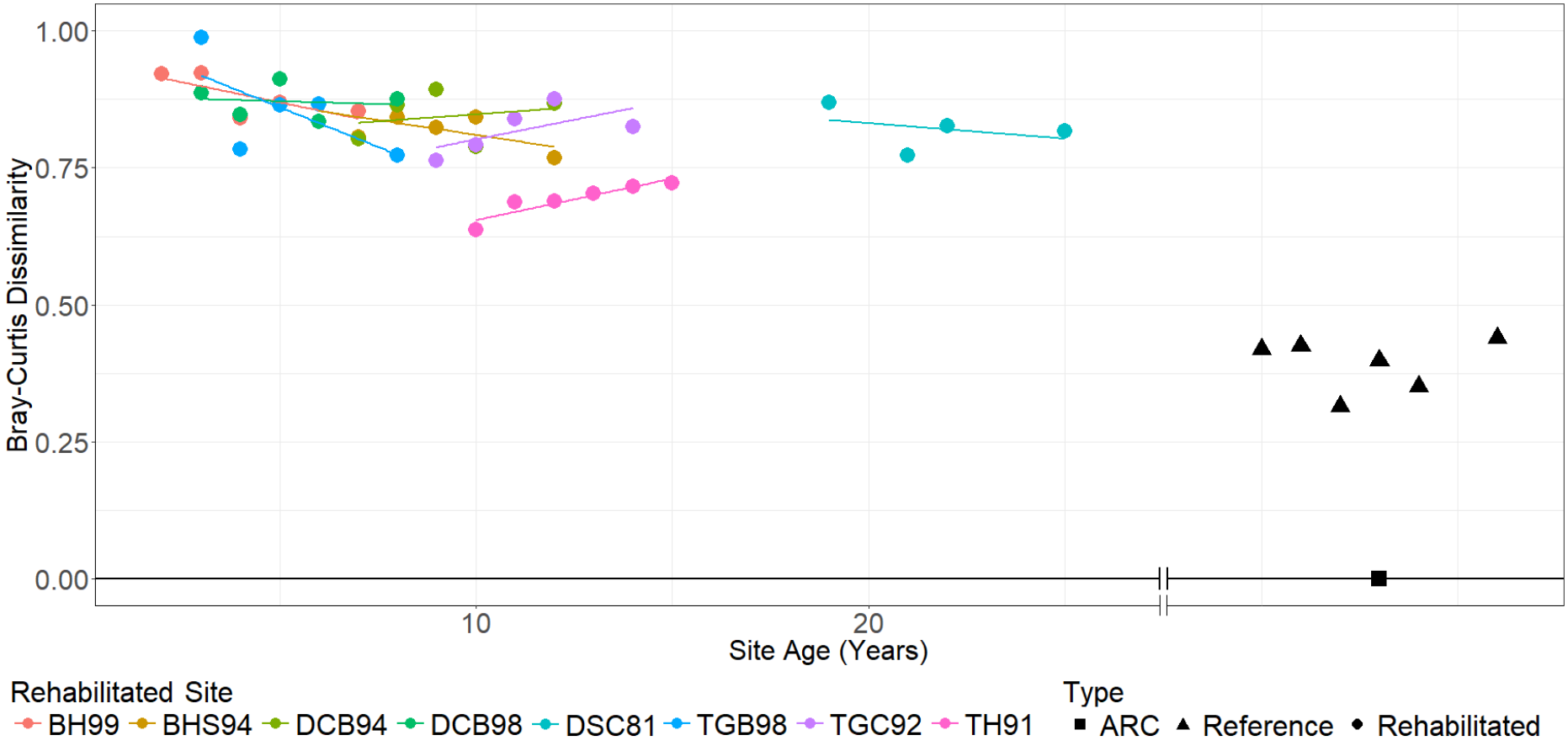
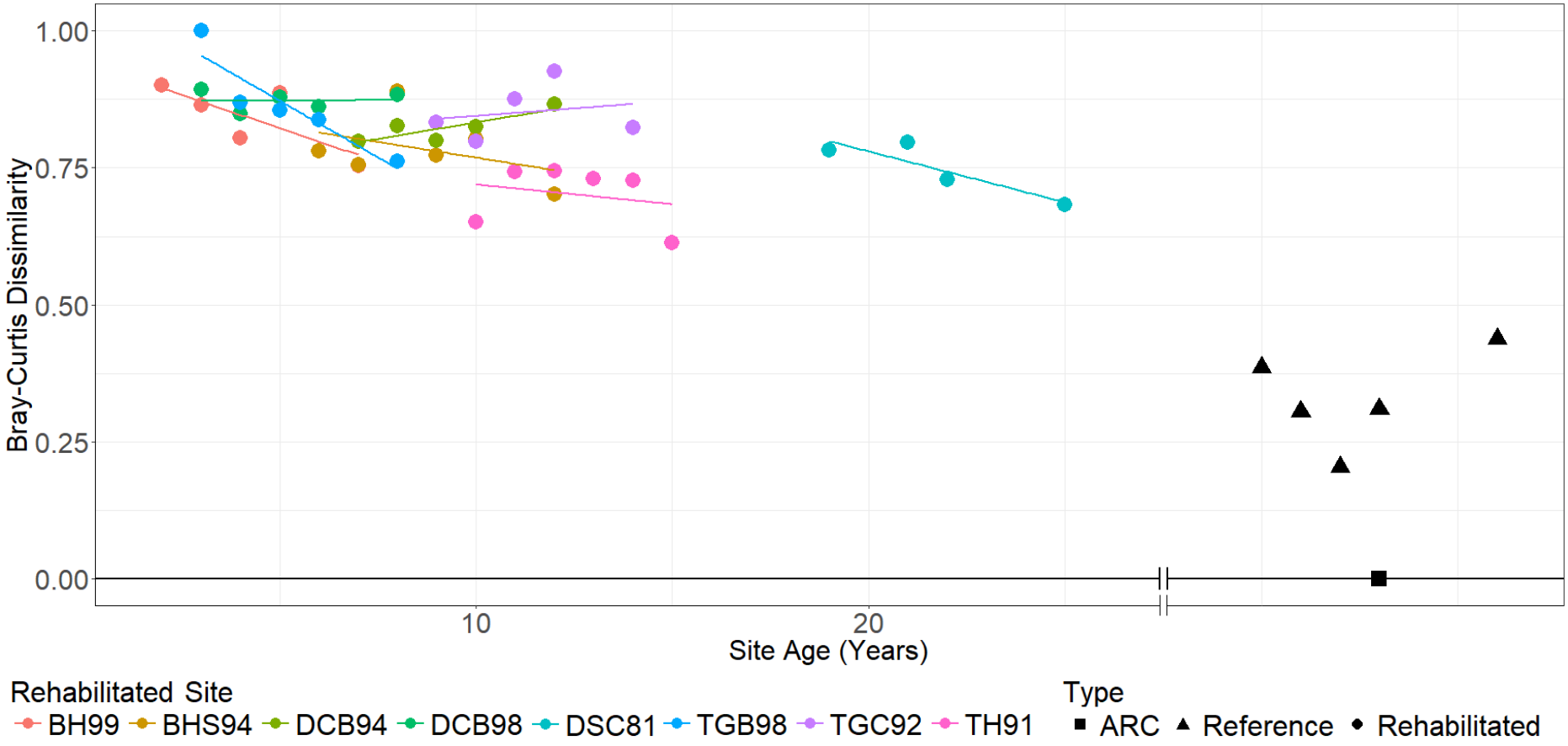


Fig 2c: *Convergence with Site 9, Callide Mine*



1.3 – Functional Group Profiles

The functional group profiles of the rehabilitated sites (figs. 24-31 (Appendix 1)) show varying levels of convergence with the reference sites (figs. 21-23 (Appendix 1)). The rehabilitated sites are all dominated by Dominant Dolichoderines, at least initially, with the exception of TH91, and Dominant Dolichoderines remain the most abundant functional group at four of the eight rehabilitated sites. However, there is some evidence of change. Rehabilitated sites BH99 (fig. 24 (Appendix 1) and BHS94 (fig. 25 (Appendix 1)) in particular show increasing relative abundances of Generalised Myrmicines and Hot Climate Specialists over the course of the sampling period, so that by the last year of sampling they had functional profiles that closely resemble the ARC profile of Reference Site 8, where over 50% of ants sampled at the site are Generalised Myrmicines. TGC92 (fig. 30 (Appendix 1)) did not achieve quite the same level of convergence, but goes from having a species composition of over 90% Dominant Dolichoderines to one of less than 50% Dominant Dolichoderines, and experienced an increase in Opportunists from minimal presence to 25% of ants sampled on site, greatly increasing its similarity to the Opportunist-dominated reference site 6 (fig. 21 (Appendix)). Contrastingly, the divergent nature of DCB94 and TH91 also appears to be driven by changes in functional group composition. TH91 moved from a functional group profile similar to reference sites 8 & 9 at the start of the study to becoming increasingly dominated by Dominant Dolichoderines while Generalised Myrmicines and other functional groups become very rare, coming to resemble the other rehabilitated sites' functional group profiles as they were at the beginning of the study period. DCB94 likewise became increasingly dominated by Dominant Dolichoderines. Overall, the changing relative abundance and dominance of Dominant Dolichoderines, as opposed to other functional groups, appeared the main factor in changing dissimilarity of rehabilitated sites from the reference sites.

2 – German Creek Mine

2.1 – Overview

At German Creek, reference and rehabilitated sites are clearly separated in NMDS (fig. 3), with evident sorting of reference sites to the left and rehabilitated sites to the right, with axis 1 being the axis of separation. Reference communities do not appear to be varying systematically over time (Andersen & Hoffmann, 2003b), meeting the primary assumption of Reference-Comparison rehabilitation assessment, with the possible exception of reference site 3. Reference site 3 appears to be systematically moving down axis 2, although given all three reference sites are prone to substantial variation in position this may just be a coincidence. Unlike at Callide Mine, the reference sites are quite separated from each other, with reference 1 separated from both sites 3 and 7 on axis 1, while sites 3 & 7 overlap on axis 1 but are separated on axis 2.

There is only one overlap between a reference and rehabilitated site on axis 1, and that is the overlap of reference 7 with site 8. Reference 7 is extremely variable in its community composition and overlaps with site 8 on both axes.

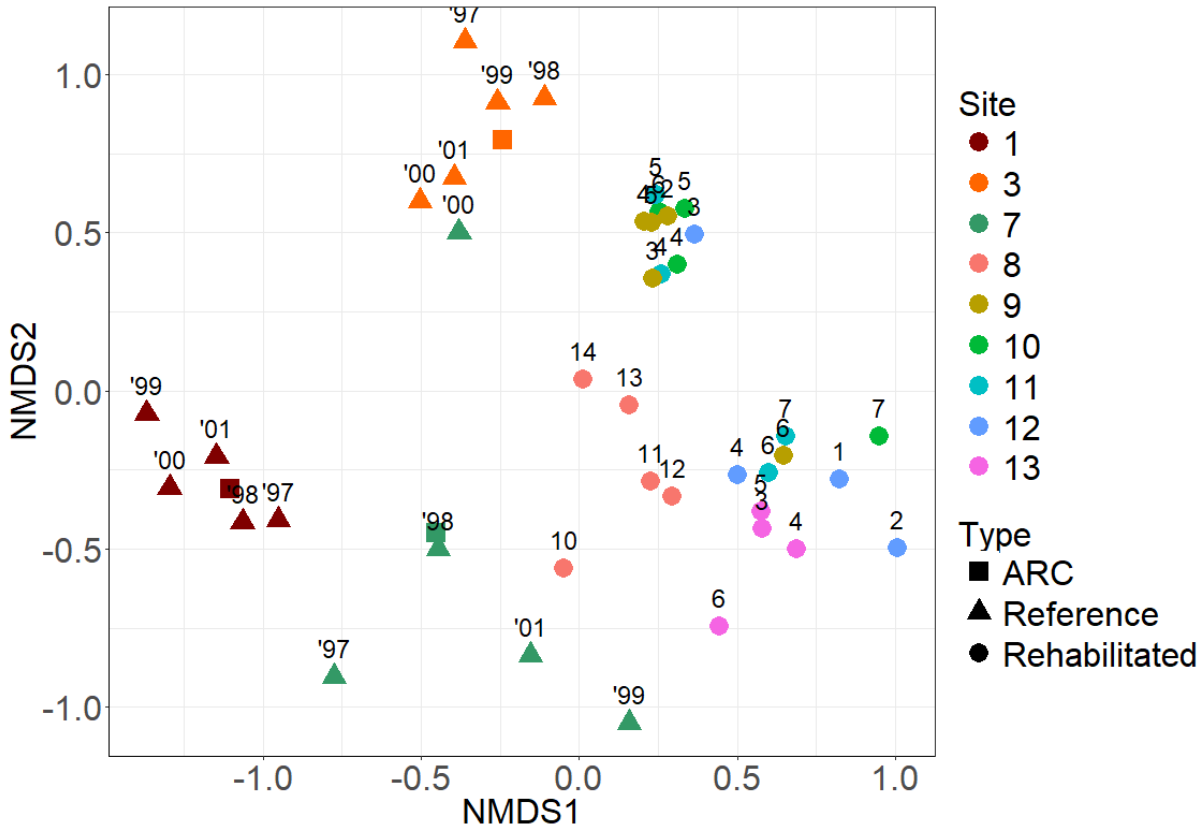


Figure 3: *Non-metric multidimensional scaling ordination of German Creek Mine Sites based on ant species abundances, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (1997-2001) for reference site samples. Two-dimensional stress = 0.197346*

In ARC-Dissimilarity (fig. 4), only site 8 is converging on **Reference Site 1**, and it is only very slowly doing so, and all rehabilitated sites are highly dissimilar to reference site 1. This reflects reference site 1 being the most far-removed of the three reference sites from the rehabilitated sites in NMDS, with no overlap on axis 1, and so the distance of rehabilitated sites to reference site 1 is largely unaffected by the movement on axis 2 that makes up the majority of rehabilitated site movement and convergence at German Creek. **Reference Site 3** shows a more conventional pattern in the sense that three of the six rehabilitated sites are converging on it, but the three diverging sites are the ones that come closest to the ARC at the beginning of the sampling period before diverging substantially by the end of the sampling period. **Reference Site 7**, uniquely among all reference sites, actually has a rehabilitated site "reach" the degree of dissimilarity from the ARC displayed by the reference site's own samples, but this is more to do with reference site 7 being unusually variable in its community composition, with correspondingly unusually high BC-Dissimilarities from the ARC.

Rehabilitated sites as a group do not significantly converge on reference sites 1 or 3 with age, but do significantly converge on reference site 7 (Site 1: slope = -0.0013, SE = 0.001, $t(19) = -0.999$, $p = 0.331$) (Site 3: slope = 0.0067, SE = 0.012, $t(19) = 0.56$, $p = 0.579$) (Site 7: slope = -0.0111, SE = 0.003, $t(19) = -3.63$, $p = 0.002$). There is substantial between-site variation in size and direction of slope relative to the reference sites, particularly in relation to reference site 3, and substantial variation in individual rehabilitated sites' slopes relative to the different reference sites.

2.2 – Rehabilitated Sites

Only one of the six rehabilitated sites at German Creek (site 8) converged with all three reference sites, possibly because, as shown in fig. 3, the reference sites differ substantially. Nonetheless, the reference sites all group on one side of axis 1, and no rehabilitated site shows straightforward convergent progression across

axis 1 towards the reference sites. Sites 12 and 13 both finish the sampling period closer to the reference sites than they started but neither move far, and site 8 makes substantial convergent progress but between the first and second years of sampling undergoes an equally large divergence. As such, the majority of convergence or divergence from specific reference sites is a result of movement on axis 2 rather than general convergence on the reference sites as a whole.

Site 8 is converging with all three reference sites (fig. 4), only marginally in the case of reference site 1 but at a brisk pace towards sites 3 & 7. These results are reflective of NMDS results (fig. 3), which shows site 8 moving systematically up axis 2 towards reference 3 throughout the sampling period, and overlapping with reference 7 on both axes, to the point of some samples from site 8 being closer to the reference 7 ARC than some of the reference 7 samples, as in the ARC-Dissimilarity results. In contrast, site 8's position on axis 1 ultimately doesn't change much and hence its relationship to reference site 1, which is located centrally in NMDS, doesn't change much either.

Site 13 remains relatively static with respect to all three reference sites (fig. 4), slightly diverging from reference sites 1 & 7, and slightly converging on reference site 3. This overall pattern is reflected in the NMDS (fig 3) as site 13 does not vary very much except for in the final year of sampling, where its movement towards the reference sites on axis 1 is largely balanced by its movement downwards away from their ARCs. What is not reflected in NMDS is the relatively consistent, if slow, convergence of site 13 towards reference site 3. Once again, NMDS is an imperfect, 2D visualisation of the data, and it is possible that higher-dimensional visualisations of this data may resolve this apparent contradiction.

In the NMDS ordination (fig. 3), Sites **9**, **10**, **11** & **12** all follow very similar trajectories and group together through most of the sampling period. They form a tight cluster with low inter-year variation throughout most of the sampling programme before diverging in down and to the right in ordination space towards the end, although site 12 begins outside the cluster and enters it only in

year 3, and site 11 splits off a year earlier than the others. This is reflected in the ARC-Dissimilarity model for **Reference 3** (fig. 4b), where all four sites show either convergent or static trends with respect to the ARC before suddenly and dramatically diverging, overriding the classification of their trends in the linear model so they are considered divergent. In the case of site 12, its initial position in NMDS (fig. 3) is actually further away than its final position post-cluster-divergence, since it began the sampling programme outside the cluster and only enters during year 3 of sampling, hence why it is still classed convergent with sites 3 & 7 by the linear model, although not with reference site 1. Additionally, moving into the cluster actually takes site 12 much closer to reference site 3, resulting in a strong convergence trend relative to reference site 3. For the other three rehabilitated sites (9, 10 & 11) however, this sudden movement out of the cluster in a divergent direction completely overrides all previous trends in the reference site 3 ARC-Dissimilarity model, while this sudden divergence is much less apparent in the ARC-Dissimilarity model of dissimilarity relative to reference site 7.

Overall the results of the German Creek rehabilitation programme are poor. No rehabilitated sites show clear and substantial convergent trajectories, with little movement by any site along the primary axis of separation between reference and rehabilitated sites. Two of the six rehabilitated sites (10 & 11) show systematically divergent trends in both NMDS ordination and ARC-Dissimilarity. The remaining four (8, 9, 12 & 13) also diverge substantially at some stage, and so do not make much progress. Despite these issues, based on the ARC-Dissimilarity and the later sampling years in NMDS, site 8 appears to be converging on all three reference sites, although this may be a product of idiosyncrasies in the relative community compositions of rehabilitated site 8 and reference sites 3 & 7. Site 12 likewise appears to ultimately be convergent with two of the reference sites. Sites 13 and 9 are converging on reference sites 3 & 7 respectively in ARC-Dissimilarity, though do not appear to be converging on these sites in the NMDS ordination, and overall there appears to be slightly less agreement between ARC-Dissimilarity and NMDS results at German Creek than

at Callide mine. That four of the sites followed divergent but very similar trends in NMDS is of ecological interest, as this suggests a distinct underlying pattern which may represent an alternative succession pathway leading to a different climax community to those represented by reference sites.

Figure 4 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 1 (4a), 3 (4b) and 7 (4c) with site age since rehabilitation (years) at German Creek Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 4a: *Convergence with Site 1, German Creek Mine*

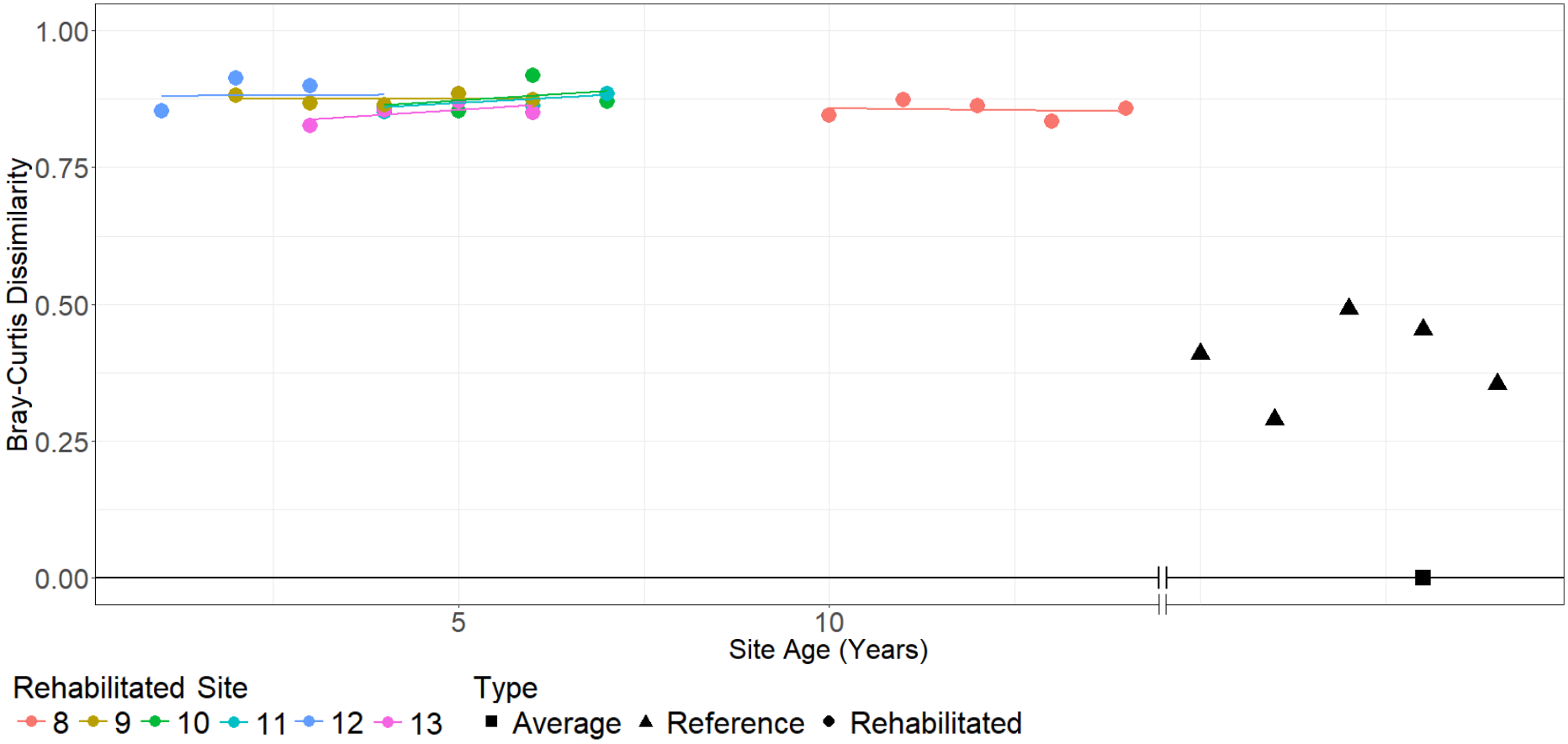


Fig 4b: *Convergence with Site 3, German Creek Mine*

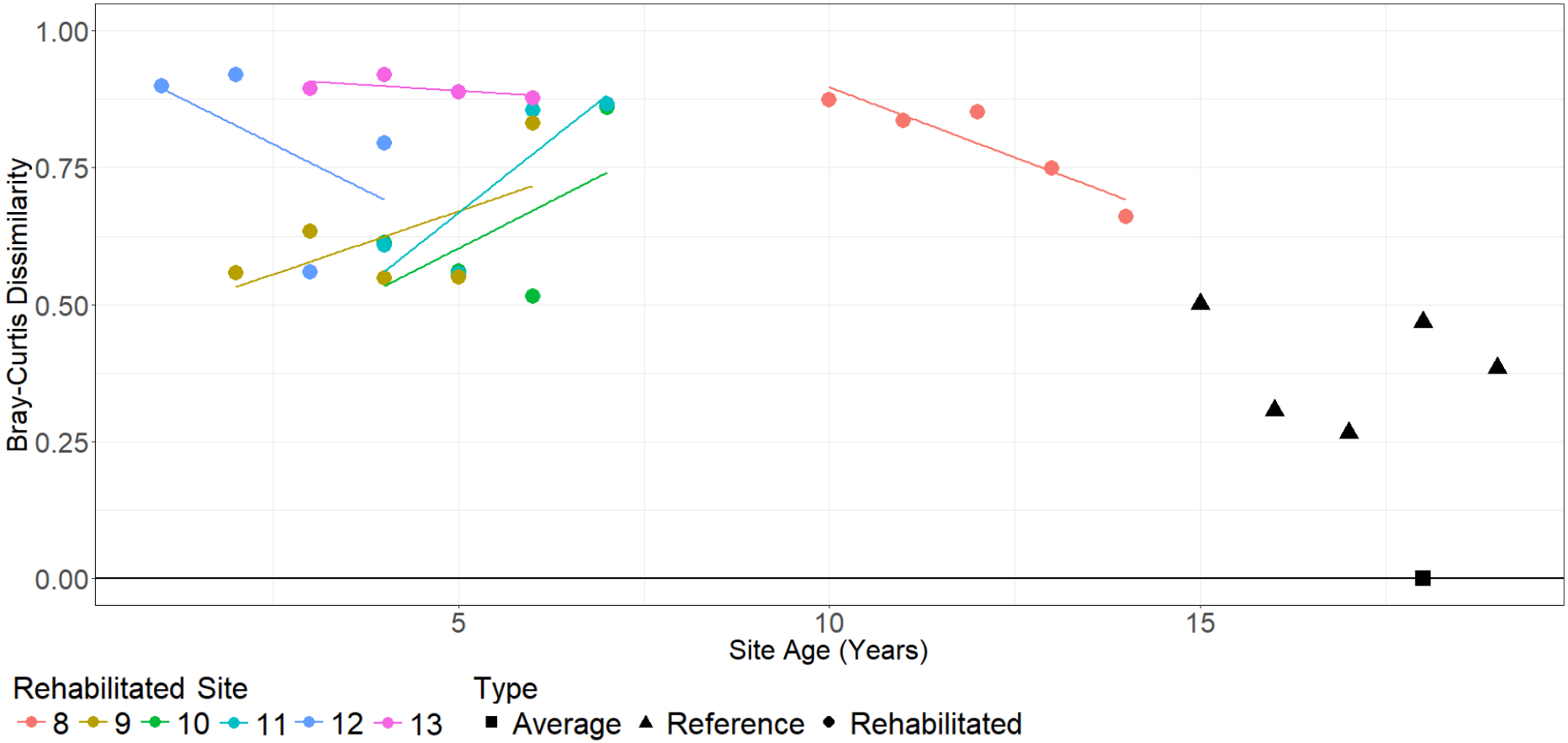
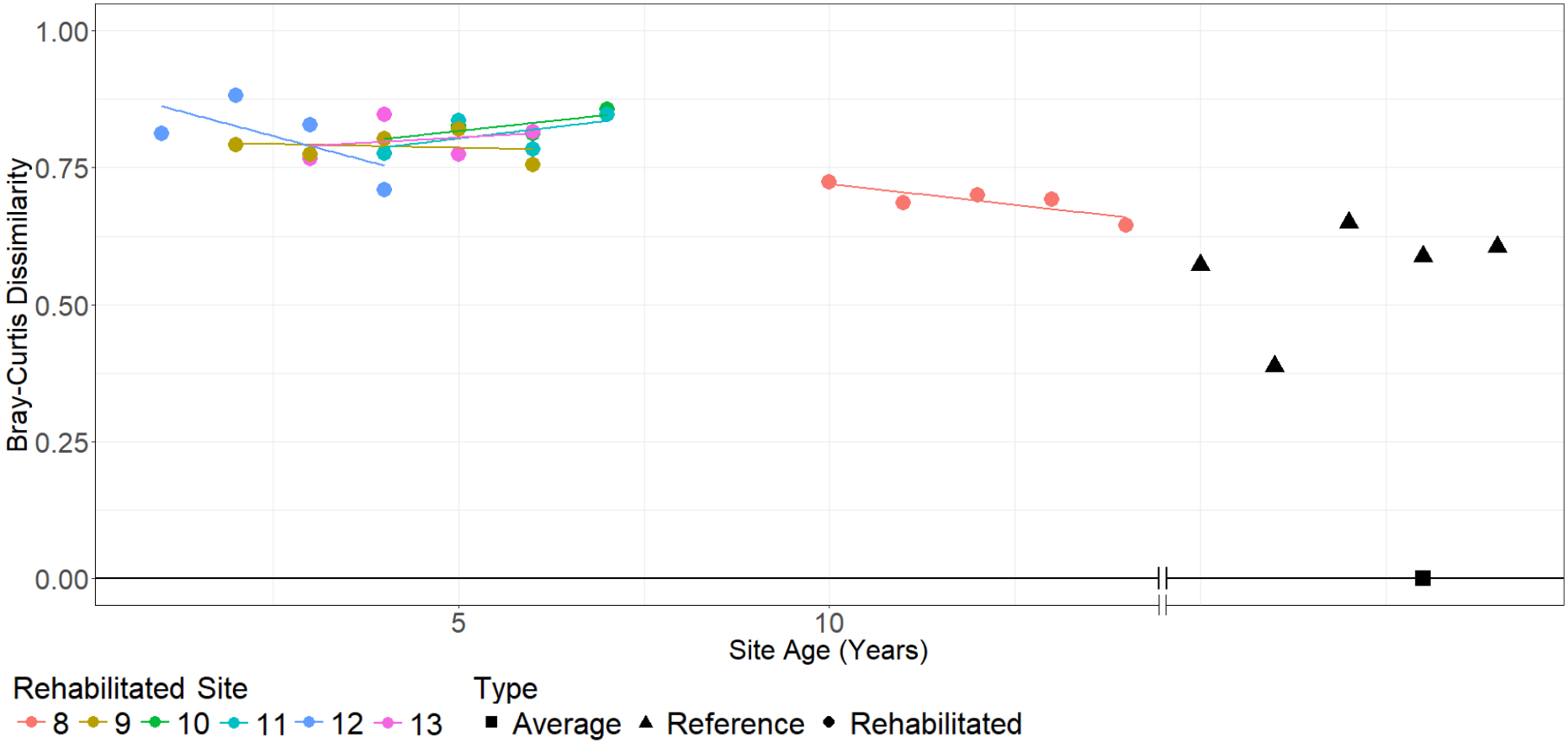


Fig 4c: *Convergence with Site 7, German Creek Mine*



2.3 – Functional Group Profiles

The functional group profiles of the rehabilitated sites (figs. 35-40 (Appendix 1)) are frequently relatively static, becoming neither more or less like those of the reference sites (figs. 32-34 (Appendix 1)). All sites are dominated by Dominant Dolichoderines, with the exception of reference site 1, where Generalised Myrmicines were more dominant in the later 3/5 years of sampling. Rehabilitated site 8, despite converging on all three reference sites, actually became gradually less like the reference sites over time with an increasing relative abundance of Hot Climate Specialists and Opportunists, before their numbers fell back down to their original levels in the last year of sampling. However, in this same year Generalised Myrmicines, present in all three reference sites at German Creek and at rehabilitated site 8 up until this time, are completely absent from the site. With the exception of site 12, which begun the sampling period with a large relative abundance of Hot Climate Specialists before dropping down to levels similar to other sites, and site 10, which experienced a large increase in the relative abundance of Opportunists in the final year of sampling, the other sites did not experience drastic changes in functional group composition during the sampling period. Changes in relative abundance of functional groups did not appear not a major factor in the changing dissimilarity of rehabilitated sites from the reference sites, with most sites only experiencing small changes in the relative abundances of each functional group.

Discussion

1 - Evaluation of Survey Methods

The information obtained about ant communities from surveys is strongly influenced by the methods used to carry out those surveys (Bestelmeyer *et al*, 2000, Longino, 2000). As such, the methods employed here warrant further examination.

1.1 – Pitfall Trapping

Ants were sampled at Callide and German Creek mines by means of pitfall traps. Pitfall traps have been found to be representative of epigaeic ant populations, with Andersen (1991, cited in Bestelmeyer *et al*, 2000) finding that results from pitfall trapping to be comparable to those obtained from intensive but relatively unbiased quadrat sampling. However, this does not mean pitfall trapping is free of bias. Pitfall trapping yield has been shown to be influenced by a variety of aspects of ant and arthropod biology. Greenslade (1973) and Andersen (1983) found that fast-moving species such as those in the genus *Melophorus* and some members of *Iridomyrmex* were more prone to falling into traps on foraging dashes, while slower ant species were more likely to either avoid or successfully re-exit traps, and Greenslade (1964) found that species of Carabidae beetles that cover more ground during foraging are more susceptible to capture, and suggested the same applies to ants (Greenslade, 1973). Greenslade also notes in both studies that pitfall trap sampling was also affected by how easy it was for arthropods to move around in the area around the trap – greater ease of movement increasing likelihood of encountering and becoming trapped in a trap. Marsh (1984) also found ant foraging patterns among species to be quite variable, with some exhibiting random foraging, some showing highly directional foraging, and also differing degrees of recruitment, contributing to variation in likelihood and evenness of trapping among species. Marsh (1984) also noted that larger species were less likely to be caught in traps they encountered as they tended to maintain contact with the trap lip and only partially enter the trap,

rather than fully entering the trap and then attempting to exit, and Olson (1991) found that the average body size of ants caught in pitfall traps was 1mm smaller than those sampled in Winkler litter sifting, though also noted that pitfall traps were better for large scavenging and predaceous ant species such as Ponerines. Perhaps most importantly for this study of mine-site rehabilitation, Majer (1997) and Olson (1991) found that pitfall trapping catches much less than the full complement of species at a site. Critically, this shortfall in species increased with habitat complexity (Majer, 1997). There are two major implications of this. The first is that reference sites, with undisturbed, complex habitats, are likely to be relatively poorly-represented compared to rehabilitated sites, leading in turn to overestimation of site "recovery" in the convergence model as pitfall trap catches potentially exclude a substantial proportion of the reference site community. The second is that as the full ant community of recovering rehabilitated sites will also become increasingly underrepresented as time goes on. This underrepresentation predominantly affects non-epigaeic species, such as cryptic, arboreal and hypogaiec species, excluding entire functional guilds of ants, in addition to rare species that are particularly prone to under-sampling (Olson, 1991, Majer, 1997).

Despite these limitations, Andersen (1983, 1991, cited in Bestelmeyer *et al*, 2000) found pitfall trapping to be equally effective as more intensive quadrat surveys, and Majer (1997) concludes that pitfall trapping remains a practical and effective means of surveying epigaeic ant fauna. Indeed, while pitfall trapping systematically excludes non-epigaeic fauna, Nakamura, Proctor & Catterall (2003) suggest that using an alternative sampling method, leaf-litter sampling, to address this exclusion results in the sampling of fewer epigaeic species, making the samples less effective for discriminating between open and forested habitats, such as rehabilitated and reference sites. These conclusions echo those of Olson (1991) who found low overlap of species between Winkler litter sampling and pitfall traps. Ultimately, many authors recommend using a suite of complementary approaches to address the biases of individual sampling methods (Olson, 1991, Majer, 1997, Nakamura, Proctor & Catterall, 2003). But

this is not always possible, and pitfall traps offer a relatively good trade-off between utility and the required effort and price required to use them, since they are cheap, relatively easy to transport, deployable for days at a time, low-tech and require little active searching effort in return for a moderately representative sample of the whole ant community and a very representative sample of the epigaeic ant community (Olson, 1991, Majer, 1997, Bestelmeyer *et al*, 2000). As such, pitfall data is a valuable tool for evaluating restoration of rehabilitated mine-sites, but we must remain aware that, like all sampling, it is an imperfect construction of reality that balances accuracy and real-world constraints on data-collection, and ideally further investigation would make use of a variety of survey techniques in order to establish a fuller picture of restoration.

1.2 – Survey Timing

The other major aspect of the survey methods due a critical eye is the survey time period. The relatively intensive sampling of the ant community, annually for 4-6 years, across rehabilitated sites of a range of ages (1-25 years since rehabilitation), combines real-time (e.g. Majer & Nichols, 1998) and chronosequencing (e.g. Bisevac & Majer, 1999) approaches for an in-depth and extensive look at changes in ant community over time. However, this is not without its flaws. The various rehabilitated sites start with a variety of initial conditions, and how each was rehabilitated is unknown and not necessarily the same for each site, given that the range of rehabilitation starting times covers almost 20 years of gradual improvements in rehabilitation technique in Australia (Bisevac & Majer, 1999, Majer *et al*, 2013). This limits what can be said about long-term trends at the two mines, as the rehabilitated sites have differing ecological histories and, as can be seen in the results, differing subsequent patterns of community change, so while some rehabilitated sites (e.g. sites 9-12 at German Creek) may display very similar trends, we cannot assume they will later come to follow the older rehabilitated site no. 13's trajectory once they are the same age as that site, and indeed for this particular pairing it is highly unlikely that this will be the case. At Callide Mine, the idiosyncrasies of the oldest

rehabilitated site, DSC81, present similar problems – its trajectory does not match up with the younger rehabilitated sites. In fact, those that are converging appearing to be doing so at a faster rate than DSC81, possibly as a result of improved rehabilitation techniques. But because we have no other rehabilitated sites of a similar age, we cannot make reliable predictions about how the communities of the younger converging rehabilitated sites will change when they reach a similar age to DSC81.

We are then largely limited to the 4-6-year scope of the surveys to evaluate the community recovery of the rehabilitated sites and must evaluate each individually. While we do make use of all-site linear models to assess changes in reference site-rehabilitated site dissimilarity, and at Callide Mine they appear to show recovery of the sites, these results should be treated with caution as they are a composite of eight individual rehabilitated sites on a variety of trajectories, varying widely in slope and even whether they are converging or diverging from the reference sites. Therefore, these overall results may not necessarily be particularly representative of individual sites' recovery and should not be treated as such. The evaluation of sites on an individual basis is still very valuable however, and indeed given the relatively rapid rates of convergence or divergence presented, particularly at Callide Mine, may well be sufficient to judge whether a site is successfully converging or not, the primary goal of this analysis, but still presents problems. The predominant issue is that, when each rehabilitated site must be assessed separately, 4-6 years of annual sampling translates to only 4-6 data points on which to perform an analysis and make an assessment. This substantially limits statistical power when assessing trends, and so all assessment must be done qualitatively, and results treated with the requisite caution. The use of qualitative assessment creates a serious risk of overinterpretation of results, particularly given the nature of ecological data, which is very sensitive to any number of factors. These can be external effects on the system such as abnormally extreme weather events such as El Niño, or even more mundane weather-related issues such as heavy rainfall during the sampling period, which at Callide Mine in 2004 supressed pitfall catches across the mine

and rendered data from DSC81 and reference site 9 from that year unusable (Hoffmann & Andersen, 2004). They can also be internal effects, such as natural cycle of boom and bust within species populations, or fluctuation of species presence and abundance during the recovery process (Kaspari & Majer, 2000). If the sampling period coincides with part, but not the whole, of such fluctuating in community composition, a very misleading picture of restoration progress, or even of the demographic stability of the natural reference communities, may be presented. The challenge of differentiating environmental noise from the true signal of convergence or divergence even presents the possibility that sites may in fact be exhibiting simple random variation in community structure, rather than systematic changes as a result of convergence with reference site communities. Fortunately, here at least we can make reliable use of strength in numbers – that many of the sites show strongly directional trends towards, or occasionally away from, the reference sites, in spite of the challenges above and variety in their origins, suggest they are in fact undergoing systematic community change rather than random drift or fluctuation, even if the exact details cannot be determined with certainty. Ideally further studies of these mines and others would be allowed to run on for decades in order to discern and verify the true trends, but such work is constrained by practical concerns of time and cost. In order to address the risk of overestimation, in this study we fitted linear models for each site, and used the (non-significant) slope of these models as a guide to whether sites should be considered convergent or not.

Another aspect of survey timing is when to sample. Numerous studies have found evidence of seasonality in Australian ant communities (Whitford & Ettershank, 1975, Briese & Macauley, 1980, Andersen, 1983, Majer, 1985, Andersen & Hoffmann, 1998a, Postle & Philips, n.d.). Although exact peak periods vary between locations and species, ant abundance and activity tend to peak in the summer months and fall to low or even no activity in the winter, particularly for seed-harvesting specialist species (Whitford & Ettershank, 1975, Briese & Macauley, 1980, Andersen, 1983, Majer, 1985, Postle & Philips, n.d., Andersen & Hoffmann, 1998a). Not all studies have followed this pattern

however, with overall ant abundance peaking in late Autumn in northern Western Australia (Postle & Philips, n.d.) and no seasonal effect being found in northern Queensland (Woinarski, 2002). Some studies in southern Australia have found that seasonal changes in the ant community also involve high species turnover, with species abundant in summer being entirely absent in winter and replaced by another set of abundant species (Andersen, 1983, Postle & Philips, n.d.), suggesting that exclusively sampling in Summer, at the peak of ant activity, may be missing out an important aspect of the ant community. In restoration studies such as at Callide Mine and German Creek Mine, neglecting these "off-season" specialists may mean missing out on monitoring for the restoration an important part of ecosystem functioning, given the importance of ants to ecosystem processes, particularly in Australia (Andersen, 1990). During the pilot survey for German Creek Mine in 1997, sampling occurred not only mid-wet season (summer, January), but also at the end of the wet season (April) and late in the dry season (winter, September) (Andersen & Hoffmann, 1998a). It was found that species richness and abundance was highest during the January sampling period, as expected from other studies, but found that functional group composition and community structure was fairly consistent through time, indicating that there was no annual turnover of species, unlike in the temperate south (Andersen & Hoffmann, 1998a). As such, survey data collected in January, at the peak of ant species abundance and richness, is in all likelihood sufficient for representative sampling of ant community composition change at German Creek Mine, and by extension at Callide Mine, given both are situated in relatively close proximity on a biogeographic scale, under similar broad climatic conditions.

2 – Evaluation of Analytical Methods

2.1 – NMDS Ordination

Multivariate ordination of site data is a well-established method for studies of ant community composition, including studies of mine-site restoration (Majer, 1997, Nakamura Proctor & Catterall, 2003, Ottonetti, Tucci & Santini, 2006, Andersen & Hoffmann, 2006, Wittkuhn *et al*, 2011). The benefit of being able to condense and visualise complex, multidimensional relationships between community samples are obvious, particularly when restricted to non-statistical interpretation of individual site trends. However, it is worth acknowledging its two notable shortcomings for rehabilitation assessments.

The first is that it is, by necessity, a simplification of a much more complex picture, reducing down the relationships between samples to a 2D plot, and converting absolute dissimilarities between sites to rank-ordered dissimilarities. While this is invaluable, it does mean that, by necessity, some aspects of the relationships between samples are obscured, or simply cannot be represented in only two dimensions, and that differences between sites cannot be quantified from these ordinations. While not really an issue for studies looking at the difference between types of site or treatments (e.g. Wittkuhn *et al*, 2011, Nakamura, Proctor & Catterall, 2003), when trying to interpret individual sites' trajectories from newly rehabilitated to fully restored communities, this loss of information, while unavoidable, means we must be even more sceptical of the details of our conclusions regarding rehabilitated site trends.

The second issue is more theoretical, that, strictly speaking, the NMDS ordination is not optimised to assess rehabilitated site recovery. An NMDS rotated to its principle components aligns the ordination to the two axes explaining most of the variation in the ordination. In these datasets, one of the two axes that explains the most variation is usually the axis that separates reference and rehabilitated sites, the axis we are most interested in, and we can use the ordination to assess the convergence, or divergence, of rehabilitated sites from the reference sites. However, while this is a quite likely outcome, it is

not guaranteed, such as in situations where some sites have no species present, as may occur when using techniques relying on restricted species lists, or sampling of sites immediately after mining or rehabilitation (Grandin, Lenoir & Glimskär, 2013, Clarke, Somerfield & Chapman, 2006).

2.2 – ARC-Dissimilarity

This chapter also made use of a novel statistical approach, the ARC-Dissimilarity approach. Derived from the NMDS approach, this technique addresses the lack of quantification of site dissimilarities and loss of data inherent in the NMDS, allowing us to quantify the convergence with or divergence from a reference site of each rehabilitated site through changes in Bray-Curtis dissimilarity. However, this approach brings with it its own assumptions and flaws. The ARC-Dissimilarity approach makes a few key assumptions about the community data that are worth spelling out:

1) That the "ARC" is a representative measure of community
2) That variance in reference site communities is natural fluctuation, but rehabilitated sites are undergoing systematic community change
3) That convergence and divergence are one-dimensional.

The first two assumptions are intertwined. The ARC-Dissimilarity model assumes that reference sites are stable communities, which may undergo natural fluctuations in ant species populations over time but in ways that are ultimately random and do not represent systematic changes in community composition (Kaspari & Majer, 2000). This assumption has already been addressed above in section 1.2 with regards to rehabilitated sites, and in practice, the assumption of non-directional shift in reference sites is generally a safe one. After all, these sites are, in theory, undisturbed sites which should not be undergoing detectable community shifts, although against a background of nearby anthropogenic disturbance from mining and broader climatic shifts as a result of anthropogenic climate change we should be wary of making such assumptions. Fortunately, our results suggest that this assumption is sound without our datasets: there is no evidence of directional change occurring at the reference sites, which are generally quite stable in both NMDS ordination (figs. 1 & 3) and in average BC

dissimilarity from the ARC (figs. 2 & 4), with no suggestion of systematic community change.

The ARC-Dissimilarity model takes this assumption to the logical conclusion that, if reference sites are stable, then in multidimensional community space, where each sample's position is defined by its abundances of each of the hundreds of species at each site, there must be a central, average point made up of the average abundances of each of those species, a stable central point around which the reference site samples ultimately display a central tendency or "orbit" as the abundances of those species fluctuate between highs and lows. That is not to claim that said point is necessarily a possible real community composition for that reference site, as species relationships may not necessarily be so linear, but that it should exist as a mathematical entity. We cannot know for sure the exact composition of this central point, as with any true value, but can calculate an approximation by finding the average of our dataset, which becomes the ARC. As the dataset is very limited, captured at an unknown point during natural fluctuations and oscillations of unknown duration, with potentially non-linear population dynamics, the median was chosen as the measure of the "average" community composition, rather than the mean; the median is insensitive to extremes in abundance, particularly species with sudden or short-term changes in abundance. However, this means that the median is therefore insensitive to groups of closely related species where each particular species may only abundant within a single year's sampling, and may not be sampled at all in other years, but as a collective they have a continuous presence and hence a notable contribution to the ecosystem that should be taken into account. This problem is particularly noticeable for reference site 7 at German Creek. This reference site plays host to 19 species of *Iridomyrmex*, but they rapidly displace each other, so only three of the species make it into the median community composition, and do so only at low "background" abundances, rather than when hyperabundant as is typical for *Iridomyrmex* species (Greenslade, 1976, Andersen, 1995). This means that even though the median abundance of the genus *Iridomyrmex* at reference site 7 is 581, the apparent median abundance is only 43, which not

representative of natural *Iridomyrmex* abundance at the site and could therefore result in inaccurate and misleading ARC-Dissimilarity convergence patterns with regards to reference site 7, as *Iridomyrmex* have minimal influence on dissimilarity scores, despite their collective importance. Under ideal conditions, where there has been extensive study of the reference sites of a study and data have been collected for many years, and their community's changes and dynamics are well-understood and verified as oscillatory rather than systematic, the mean would be a more appropriate measure of the central point - but in these shorter-term studies the median remains a more appropriate measure of the central point of a reference site community, and we must assume variation is non-directional.

The more practically challenging assumption of the ARC-Dissimilarity approach is the third assumption, of one-dimensionality. ARC-Dissimilarity measures dissimilarity in a single dimension, the Bray-Curtis dissimilarity between a sample and the ARC of a reference site, while NMDS shows a multidimensional pattern, albeit simplified down to two dimensions. ARC-Dissimilarity is built around the goal of obtaining the true dissimilarity between a rehabilitated site and the ARC, and how that changes over time. To achieve this end, all other data is discarded, and with it potentially important contextual information. This includes loss of the trajectory a site is taking through community space. A shallow convergence slope in ARC-Dissimilarity may mean a rehabilitated site is slowly but directly heading for the reference site ARC, taking a very erratic path towards the reference site that nonetheless has a net convergent effect, or it may mean that it is on a trajectory that is not convergent with the reference site at all but happens to be moving closer in its passage to another location in community space.

This is particularly problematic in the case of sites such as DSC81 or TGC92, which appear to be rapidly converging with reference site 6 in ARC-Dissimilarity (fig. 2a) when NMDS (fig. 1) suggests that they are in fact on trajectories that will probably take them on past reference site 6 onto divergent trajectories. That said, this is not necessarily an inherent problem of ARC-Dissimilarity, but one of

the small size of the dataset for each site. This is best demonstrated by DSC81. Looking at the ARC-Dissimilarity results for DSC81 converging with reference site 6, we see that although the fitted slope is convergent, the sample taken from the last year of sampling actually shows a large increase in divergence from the previous sampling time point, coinciding with the point in the NMDS ordination when DSC81 appeared to bypass reference site 6 and begin diverging. Though we should refrain from overinterpreting such a small amount of data, it does highlight that in the event that surveying was continued for a longer period, so that the changes in community composition beyond this time point could be seen, and more sophisticated models could be fitted than the highly reductive linear models used here out of necessity, there is no inherent reason the ARC-Dissimilarity approach would not be able to pick up on the eventual divergence of sites which are not truly convergent. However, given that there have only been a handful of studies of mine rehabilitation that have been run for longer than this one as of the time of writing (Andersen & Majer, 2004, e.g. Majer & Nichols, 1998, Majer *et al*, 2013), this problem is likely to be a persistent problem. As such, it is recommended that ARC-Dissimilarity is used in conjunction with NMDS ordination, as the two approaches are complementary in their differing approaches to simplifying complex community data, and cover each other's weaknesses well, with ARC-Dissimilarity drawing out and quantifying relationships between reference and rehabilitated sites indicated by NMDS, and NMDS providing vital context to these relationships missing from ARC-Dissimilarity outputs. Co-deployment of these approaches is facilitated by their use of the same data inputs and initial calculation steps, diverging only after computation of the dissimilarity matrix.

2.3 – The Reference-Site Convergence Model

Finally, we turn to the underlying model of both approaches to mine site rehabilitation assessment, the reference-site convergence model. This simple premise – comparing rehabilitated site communities to those of local undisturbed ecosystems to see how they compare to what we would expect under natural conditions we are aiming to restore – makes for a very intuitive

and useful model, without which the task of assessing whether sites are being "restored" would be much more difficult. However, by defining restoration as "becoming more like the reference sites", an assessment excludes alternative restoration possibilities. In the first instance, rehabilitated sites that are not converging on the reference sites selected may still be converging on other local ecosystems that are not represented by the reference sites. But they may also be on a completely different, but still legitimate, restoration trajectory. Post-mining rehabilitation programmes have become increasingly sophisticated, but older programmes resorted to planting of non-native plants (Bisevac & Majer, 1999, Majer *et al*, 2013). Even in the event that rehabilitation programmes have made exclusive use of local plants, mining areas have undergone major changes in topography, in water retention, in physical structure, chemical composition and nutrient content of soil, among other factors (Bell, 2001, Banning *et al*, 2008, Pandey, Agrawal & Singh, 2014), meaning their starting conditions are not the same as those of local undisturbed ecosystems, and we should not necessarily expect them to necessarily develop to closely resemble undisturbed communities, even while attaining structurally complex habitats with diverse, late-successional biota (van Hamburg *et al*, 2004). As such, while the reference-site convergence model allows us to positively confirm when a site is undergoing recovery to a local natural state, we cannot confirm that non-converging sites are not also undergoing recovery.

2.4 – Conclusions on Methods

This study makes use of a well-established study design, with the exception of the novel analytical approach to this type of data of ARC-Dissimilarity. However, said design is not without limitations, particularly when based on effectively small datasets as is the case here. This is not to suggest that the study design employed here is inappropriate, simply that it has flaws that must be taken into consideration when results are being interpreted and management policy for sites implemented based on those results.

3 – Discussion of Results

6/8 sites at Callide Mine are converging on at least one reference site, while 4/6 sites at German Creek are converging on at least one reference site. The convergence of sites at Callide Mine is happening much more rapidly and clearly, with some sites (DSC81, BSH94 & TH91) coming close to overlapping with reference site 6 in NMDS ordination (fig. 1) and all six convergent sites showing clear convergence. In contrast, at German Creek, convergence, when it is occurring, is occurring at a much more gradual rate for the most part and is not as direct as at Callide Mine.

Also of interest is the fact that where convergence did occur at German Creek, it was often not on the primary axis of convergence (fig. 3), but along the secondary axis which separated out the reference sites, rather than along the axis that separated them based on community composition qualities possessed by the reference sites but not the rehabilitated sites. Hence while rehabilitated sites did converge on the reference sites at German Creek, they did so in ways that increased their similarity to particular reference sites, rather than in ways that made them more like the reference sites as a whole. This suggests convergence was not associated with increases in similarity in terms of presence and abundance of species that were associated with undisturbed ecosystems in general, but those with site or habitat-specific species unique to each reference site. This is reflected by the functional group profiles of the rehabilitated sites at German Creek, which, compared to those at Callide Mine, are already fairly similar to the local reference sites, with broadly similar relative abundances of functional groups, generally being dominated by Dominant Dolichoderines in most years and with a variable but consistently present set of Hot Climate Specialists, Generalists and Subordinate Camponotini and an infrequent and inconsistently sampled population of Generalised Myrmicines. Hence changes in community composition at the species level were not associated with major changes in functional group composition, meaning that the differences between reference and rehabilitated sites at German Creek was not closely associated with major ecological changes in the ant community, but rather with turnover of

the species and genera within each functional group, with the broader ecological framework of the ant community remaining relatively unchanged (Andersen, 1995). That said, the reference sites at German Creek were less similar to each other than at Callide Mine, so we should perhaps not expect to see as much convergence on reference sites as a group when they did not share as much of a collective community composition to converge with. Exploration of the species associated with these convergence patterns would help shed light on the differences in convergence patterns between the two different mines.

On a related note, there was more convergence on reference sites that are more initially similar to the rehabilitated sites at both mines, with rehabilitated sites at Callide Mine appearing to be converging on reference site 6 in particular even as they converge on all reference sites, and only one rehabilitated site converging on reference site 1 (site 8, which converges on all three) at German Creek while three rehabilitated sites apiece converge on reference sites 3 and 7. Whether these sites represent dominant ecosystems in the local area or whether rehabilitated mine sites are more likely to follow successional trajectories that lead to more "mine-like" sites, is unclear from this data alone. The functional group profiles for Callide and German Creek present a mixed picture. At Callide Mine, three of the six converging sites (DSC81, TGB98 & TGC92) show rising relative abundances of Opportunists which put them on a trajectory to becoming like reference site 6, the site they appear to be converging on in NMDS and ARC-Dissimilarity, but of the remaining three sites, two (BH99 & BHS94) became dominated by Generalised Myrmicines (<50% of community), putting them on a clear trajectory for reference site 8. This pattern could then, at Callide Mine at least, simply be an artefact of the position of reference site 6 between the rehabilitated sites and the other two reference sites – rehabilitated sites converging with reference sites 8 & 9 will initially become less dissimilar to reference site 6 as they approach it before "passing by" as their community composition moves towards that of reference site 8 or 9. But at German Creek the four convergent sites all have functional group profiles similar to either reference site 3 or 7, rather than the more dissimilar reference site 1 with its

unique Generalised Myrmicines-dominated community, though this site could simply be an unusual ecosystem for the local area that mine sites would not generally develop into. The question thus remains unclear, and could potentially be resolved with further study of mine site recovery and whether sites with lower initial dissimilarity to rehabilitated sites are more likely to turn out to be end-points for rehabilitation succession.

At German Creek, the unusual clustering and shared fates of rehabilitated sites 9, 10, 11 & 12, at least according to the NMDS ordination (fig. 3), are particularly interesting in light of prior discussion of alternative restoration pathways (section 2.3). These sites were all rehabilitated within a few years of each other and show similar trends, although less so in the case of site 12. However, despite being rehabilitated at around the same time, rehabilitated site 13 follows a completely different trajectory. Comparing factors such as rehabilitation technique and environmental conditions at these sites may help tease out why they appear to have such different community compositions and fates, and in doing so give us insight into what drives post-rehabilitation succession in certain directions. One factor that can be ruled out however is functional group composition. The four clustered rehabilitated sites do not share any distinct traits in their functional group profiles or in how those profiles change over the sampling period. Nor is rehabilitated site 13 distinguishable from these four sites based on functional group profiles alone.

One of the most interesting results concerns rehabilitated site TH91 at Callide Mine. In NMDS ordination, TH91 is unique in that it groups with reference sites 8 & 9, and is relatively stable, shifting very little in comparison with other rehabilitated sites. This stability and close proximity to reference sites are features we might expect in a fully restored site. However, appearances can be deceptive, as ARC-Dissimilarity shows that TH91 is actually diverging, at an appreciable rate, from reference site 8, and shows an arc-shaped pattern relative to reference site 9. However, a closer examination of the ARC-Dissimilarity pattern of TH91 relative to reference site 6 shows relatively little net change

overall, from which we may infer that TH91 may be demonstrating the kind of stable, oscillating pattern we would expect from a reference site, rather than truly systematic movement. This site warrants further study and sampling, to unravel whether this site has truly achieved restoration. That said, the site's functional group profile also reveals that the functional group composition of TH91 changed from one similar to that of reference sites 8 & 9, with relatively high abundances of Generalised Myrmicines of around 30%, to one resembling an early-stage rehabilitated site at Callide Mine, where over 90% of ants are Dominant Dolichoderines, a fate shared with the other diverging reference site, DCB94. This marks a transition from functional group composition resembling that of a local natural community to one resembling a highly disturbed site, indicating a failure of restoration (Andersen, 1995). However, even if TH91 has not achieved true convergence and is falling into decline, the apparent rapid advancement from rehabilitation to a very reference-site-like community, even if it is not necessarily as stable as it appears, is of great interest, and it would be worth investigating the starting conditions of this site and its current dynamics to attempt to determine why this site has progressed so much more swiftly than the rest to a near-restored state, even if there is no data of the actual transition period. Of course, given prior discussion of the various limitations of these methods for determining restoration, TH91 may not be nearly so close to fully restored as it appears, but even if it was only restored in terms of epigaeic ant fauna, that is worthy of investigation in order to determine how this has come about when no other site has achieved this.

The functional group profiles of the sites at each mine reveal different demographic processes are driving convergence patterns at Callide Mine and German Creek Mine. At Callide Mine convergence and divergence is often associated with changes in the relative abundances of the functional groups, particularly in the relative abundance of Dominant Dolichoderines. Rehabilitated sites tend to be dominated by Dominant Dolichoderines, primarily species of *Iridomyrmex*, and particularly by *I.* species P (*rufoniger* gp). *I.* species P appears at incredibly high abundances, frequently in greater abundances than all other ants

in a sample and often many times greater, in 1/3 of rehabilitated site samples across 6/8 of rehabilitated sites. It does not occur in such huge abundances in any of the reference sites, with the exception of reference site 8, and even here it only occurs in large numbers in two samples and its abundance is relatively modest. Reference samples at Callide Mine tend to have lower total ant abundances and greater relative abundances of other functional groups, particularly Generalised Myrmicines, Opportunists and Hot Climate Specialists, and convergent rehabilitated sites show increasing relative abundances of these functional groups too, while those that are diverging, such as TH91 and DCB94, show the opposite trend.

In contrast, at German Creek the functional group profiles of reference and rehabilitated sites are relatively undifferentiated, all being dominated by Dominant Dolichoderines with low numbers of Hot Climate Specialists and Opportunists and very low numbers of Generalised Myrmicines & Subordinate Camponotini, with the exception of reference site 1 with its elevated percentages of Generalised Myrmicines and Opportunists. Convergence and divergence is therefore driven by turnover within the functional groups, rather than major transitions in the community ecology. This is best exemplified by the convergence and divergence of rehabilitated sites to reference site 3. At reference site 3 the high relative abundance of Dominant Dolichoderines is largely as a result of the presence of a highly abundant *I. rufoniger* species, species E. *I.* species E is the most abundant species at German Creek, dominating 11/41 site samples from reference and rehabilitated sites alike, across 6/9 sites. However, *I.* species E is not the only highly abundant *I. rufoniger* gp species present at German Creek. *I.* species B (*rufoniger* gp) is not as abundant as species E, but dominates site samples with similar frequency, and the two alternate as the dominant species at four of the six rehabilitated sites. Due to their hyperabundant nature, they exert a strong effect on Bray-Curtis dissimilarity of the rehabilitated sites from the reference sites, particularly from reference site 3, where the ant population also has a large majority of *I.* species E (*rufoniger* gp). Consequently, the biggest factor in rehabilitated sites' dissimilarity to reference

site 3 is which species of the *Iridomyrmex rufoniger* species group is currently dominant at the site, resulting in the large and sudden divergences observed in the convergence of sites 9, 10 & 11 in 2000 and 2001, as *I.* species B became dominant at sites 9 & 10 in 2001 and at site 11 in 2000, all three sites having previously been *I.* species E-dominated in line with reference site 3. Site 12 also undergoes rotation of the dominant *I. rufoniger* gp species and subsequently diverges and converges with reference site 3. This feature of the German Creek site is the cause of striking clustering of sites 9, 10, 11 & 12 and their shared trajectory at German Creek. The turnover of dominant *I. rufoniger* gp species at the rehabilitated sites at the Mine leads to all four clustered together in a position of low dissimilarity with reference site 3 because they all share the trait of having an ant community largely composed of *I.* species E with reference site 3 and each other. Rehabilitated site 12 shifts into this cluster when its dominant *I. rufoniger* species rotates from species B to species E. Then all four move away from reference site 3 to a position close to rehabilitated site 13 and to where rehabilitated site 12 was previously positioned, because their dominant *I.* rufoniger species is now species B and they share more in common with rehabilitated site 12 in years 1 & 2 and site 13, which is consistently dominated by species B. Hence the movement of these sites is movement between two types of community composition – dominated by *I.* species E (*rufoniger* gp) or *I.* species B (*rufoniger* gp) – as the two hyperabundant *Iridomyrmex* species replace each other over time. In contrast, reference site 7 is not dominated by one of *I.* species B or E in most years, instead being host to a variety of other species of Dominant Dolichoderines simultaneously, and hence convergence patterns of rehabilitated sites with this reference site are not defined by which *I.* rufoniger species is dominant at the rehabilitated sites at the time.

The results of the analysis of the success of the rehabilitation programmes at restoring the native epigaeic ant fauna at these two mine sites reveal varying levels of success between the two, but also somewhat different ecological conditions in regard to reference sites. But of greater ecological interest is the implications of the trajectories of various rehabilitated sites for restoration as a

whole, offering a tantalising glimpse into the factors that determine the successional trajectories of rehabilitation sites, and hence possibilities for improving rehabilitation procedures to better shape these trajectories.

Chapter 3 – Assessing the efficacy of four simplified analysis protocols for evaluating rehabilitation success

Introduction

Ants are as ubiquitous in Australian environmental monitoring programmes as they are in the Australian fauna. This is a direct result of ants being one of the most well-suited taxa to the task of bioindication on the continent, since they are conceptually excellent indicator components, a highly representative group for many other plants and invertebrates, and important contributors to ecosystem processes and patterns in their own right, as well as being more practically convenient to use than other invertebrates (Andersen, 1990). However, though they are the best of the options available, they are far from perfect organisms for general use as bioindicators.

Much of this difficulty comes from the sorting and identification of samples. When the species number is very high, as in large parts of Australia, and the number of samples is high, as in a large, multi-site monitoring programme such as at Callide and German Creek, this process is extremely time-consuming, even for specialists (Andersen, 1990, Bisevac & Majer, 1999, Lattke, 2000, Arcoverde *et al*, 2017). For the data analysed in this study alone, over 140,000 individual ants had to be sorted and identified to species. Furthermore, for the majority of species, true species identification was not possible, and samples instead had to be designated with a study-specific species code within a species-group, due to the poor taxonomic coverage of Australian ants relative to their diversity (Andersen, 1990). This not only prevents comparison of individual species trends between studies to help establish patterns, it also means specialist knowledge of ant taxonomy is required for every study (Andersen *et al*, 2002, Arcoverde *et al*, 2017). Even when identification to species is possible, specialist knowledge is still required to differentiate between members of the same genus (Andersen, 1990), meaning even a full taxonomy would do little to make the use of ants as

bioindicators more accessible to land managers and other non-specialist end-users (Andersen *et al*, 2002, Arcoverde *et al*, 2017). In terms of the practical details of obtaining data, the hyper-diversity and abundance that make ants such good indicator components ecologically work against them. The time and cost associated with having specialists identify each and every sample is prohibitive to the adoption of ants into monitoring programmes as standard protocol and to longer-term monitoring where they are already included.

The solution to this roadblock to the widespread uptake of ants as an indicator taxon is to find surrogate taxonomic groupings as a substitute for ant species abundance data, which are simpler to sort and identify. This would reduce time and funding required to obtain ant community composition data and use it as an indicator for evaluating rehabilitation progress. A suitable surrogate taxonomic group for use in a simplified analysis protocol must be able to replace an analysis using full assessment of the abundances of all species (Species Abundance[1] analysis) without significant changes to the analysis's evaluation of rehabilitated site recovery progress (Andersen *et al*, 2002). Four potential simplified analyses are tested in this chapter:

1) Genus Abundance[2]
2) Functional Group Abundance[2]
3) Large-Bodied Species Abundance[2]
4) Species Presence/Absence[2]

---

1. Capitalisation of Species Abundance, refers explicitly to the evaluation of rehabilitation success at Callide Mine and German Creek presented in chapter 2. Use of lower case species abundance refers to abundances of species.

2. As with capitalisation of Species Abundance, the capitalisation of Genus Abundance, Large-Bodied Abundance, Functional Group Abundance & Species Presence/Absence refer explicitly to the simplified analyses of rehabilitation success at Callide Mine and German Creek using abundances of genera, functional groups and large-bodied species and presence of species presented in Chapter 3. Use of lower case refers instead refers to abundances of genera, functional groups and large-bodied species and presence of species respectively.

The four simplified analyses selected were identified from the literature on the use of ants as bioindicators, being approaches that had been shown to work elsewhere. The use of **genera abundances** is regularly identified as a good surrogate for species abundance information across a wide variety of taxa, including ants (Pik, Oliver & Beattie, 1999, Gallego *et al*, 2012, Ribas & Padial, 2015, Souza *et al*, 2018). This idea is grounded in the phenomenon of niche conservatism, the tendency for species to "retain ancestral ecological traits and environmental distributions" (Crisp *et al*, 2009), and therefore for congenerics to share similar traits and distribution (Crisp *et al*, 2009, Ribas & Padial, 2015). As such, moving up by a single taxonomic rank is likely to result in little loss of precision regarding species responses to environmental changes, as congenerics should respond in similar ways. For ants, identifying samples only to genus level is advantageous not only as a time-saving measure, but also for accessibility – unlike species, the different genera are quite distinctive, and the identification of genera is relatively simple task with many resources available to aid with identification, that can therefore be performed even by non-specialists, and for specialists is an easily and rapidly accomplished task (Andersen, 1990, Brandão, 2000, Lattke, 2000).

Functional Group Abundance here refers to the Functional Group scheme, discussed in chapter 2 (Methods, section 3), which divides ant genera into nine functional groups based on their competitive interactions with each other, biogeography, and tolerance to temperature, stress and disturbance. Despite the emphasis on ecology and behaviour over taxonomy in this classification scheme, there is still a strong taxonomic basis for diagnostic traits, such that group affiliation is generally decided at the level of genus, demonstrating niche conservatism in a manner similar to genera (Andersen, 1995a, Crisp et al, 2009). A Functional Group Abundance simplified analysis offers similar advantages to that of Genus Abundance analysis, largely because membership of the functional groups is usually classified based on genera, meaning it is usually only necessary to identify samples to genus, which, as discussed above, is relatively simple and can even be done by non-specialists. While being slightly more difficult to apply

for non-specialists through the need to identify some samples to species and species groups, using the Functional Group Scheme also enhances the comparability of results across studies and even across continents due to its origins as a biogeographic-scale comparative tool (Andersen, 1995a), making it a powerful analytical tool in a field constrained by limited replication and cross-site applicability of results (Majer & Nichols, 1998).

Unlike the previous two proposed simplified analyses, the use of the abundances of **Large-Bodied species** reduces effort required through restricting the taxa of interest to in effect only those that are relatively easily to identify. The "Large-Bodied" protocol originates in the work of Andersen *et al* (2002), who as part of their simplified protocol considered only species belonging to genera and species groups with a total individual worker body length greater than 4mm. For polymorphic ant species, only those with minors larger than 4mm were included (B. Hoffmann, pers. comm.). Large ants are much more readily sorted to species and identified to genus than smaller ants, making them much more accessible to non-specialists (Arcoverde *et al*, 2017). The use of large ants alone is not simply a selection based on convenience, however, although that is an important benefit. Larger species have previously been shown to be representative of overall ant community responses to sulphur dioxide emissions and grazing intensity (Andersen *et al*, 2002, Arcoverde *et al*, 2017), and are particularly sensitive to disturbance (Ness *et al*, 2004, Gibb *et al*, 2018), making them a good indicator of restored ecosystems.

The final proposed simplified analysis, **Species Presence/Absence**, is, like genus abundance and functional group abundance, commonly used as a substitute for species abundance in the literature and is in fact often combined with other simplified analysis measures (Wright *et al*, 1995, Thorne, Williams & Cao, 1999, Pik, Oliver & Beattie, 1999, Andersen *et al*, 2002, Marshall, Steward & Harch, 2006, Ribas & Padial, 2015). Species Presence/Absence is binary community data which only takes whether species are present or absent in each sample into account, rather than how many individuals of each species are present. This

means that a species with only one individual in a sample is given equal weight in the analysis to a species with an abundance numbering in the hundreds or thousands. Species Presence/Absence bears a greater resemblance to Species Abundance analysis than the other proposed simplified analyses, sharing the same level of taxonomic resolution, unlike Genus or Functional Group Abundance analyses, and making use of the full complement of species, unlike Large-Bodied Abundance analysis. As such it is likely to produce the most similar results of the four to those of Species Abundance analysis. However, the trade-off for that increased likelihood of fidelity is that it saves less effort than the other analyses, as the identification of samples to species is still required. Crucially, this means that even though Species Presence/Absence analysis cuts down the time required to process samples, it still requires a specialist taxonomist, so lacks the broader accessibility of the other three analyses.

The four simplified analyses thus represent a continuum of ease of use and applicability, with Genus Abundance being the easiest to use and Species Presence/Absence being the most difficult, aside from Species Abundance itself. But the utility of simplified analyses is dependent not just on how easy they are to use, but on how closely they adhere to the results obtained from a full analysis using species abundances to measure community composition. Hence, the aim of this study is to determine how well each of the four analyses replicates the results of the Species Abundance analysis carried out in chapter 2.

Methods

1 – Simplified analyses

Datasets for the four simplified analyses were derived from the original species abundance datasets for Callide and German Creek mines (see Chapter 2 Methods for details). For Genus Abundance, species were sorted into their respective genera and their abundances pooled, and the same procedure was applied when sorting species into one of the nine functional groups, following the classification in Andersen (1990). Total abundances of each functional group were used rather than relative abundance for consistency with other analyses. The large-bodied species dataset was created by compiling the abundances of qualifying species in each sample from the full species abundance dataset, following the classification outlined in Andersen *et al* (2002). Where it was unclear from the literature what functional group a species belonged to or whether it qualified as large-bodied, the species was categorised on the basis of expert opinion (B. D. Hoffmann, pers. comm.). The species presence dataset was derived by conversion of species' abundances to presence/absence data. Data for the other three analyses were square-root transformed to control for highly abundant species. Samples excluded from the original analysis were also excluded in these simplified datasets (table 3 (Chapter 2 Methods)).

2 – Comparison with Species Abundance analysis

The viability of each of these four simplified analyses as a substitute for a full species abundance assessment of a mine site rehabilitation programme is dependent on their capacity to reproduce the key aspects of community change trends from the species abundance assessment. In the case of mine site recovery, the key aspects are the direction of travel for rehabilitated mine site communities – are they becoming more like reference sites, converging with them, suggesting a return to a natural community composition for that locality, or are they static with respect to reference sites, or even becoming less like them, diverging from them, suggesting, under the convergence model of

recovery assessment, that they are in fact not recovering to a natural community composition? A secondary aspect of this question is how closely these sites reproduce those trends, whether they are able to reproduce the rate of recovery of a rehabilitated site, telling us whether a site is on track to recovery on an acceptable timescale or will not recover for decades or centuries.

The recovery of rehabilitated mine sites at Callide and German Creek mines was initially assessed using NMDS ordination and ARC-Dissimilarity based on Bray-Curtis dissimilarity (Chapter 2), so the four simplified analyses were likewise assessed with these techniques, with reference site ARCs, Bray-Curtis dissimilarity matrices and NMDS ordinations calculated from the new datasets. In the case of Large-Bodied Species, one site sample at Callide Mine (TGB98, Y1) had no large-bodied species. In order to correct for this, a zero-adjusted Callide Large-Bodied Species Abundance dataset was used with a dummy species inserted into the dataset with an abundance of 1 at every site sample (Clarke, Somerfield & Chapman, 2006). In order to evaluate the fidelity of the results of the simplified analyses to the original species abundance assessment, four approaches were used:

- Procrustes analysis on the NMDS results (Buttigieg & Ramette, 2014)
- Cohen's Kappa Inter-Analysis Agreement on ARC-Dissimilarity Rehabilitated Site Convergence (Cohen, 1960)
- Intraclass Correlation Inter-Analysis Agreement on ARC-Dissimilarity Estimated Years to Convergence (EYCA) (McGraw & Wong, 1996)
- Qualitative interpretation of NMDS Ordination patterns

Procrustes analysis was used to compare the overall shape of NMDS ordinations for each simplified analysis to the species abundance analysis ordination (fig. 1 (Chapter 2 Results)) and rank them in order of most to least similar. The zero-adjusted Callide Large-Bodied Species Abundance NMDS was compared to a Callide Species Abundance NMDS calculated from a Callide Species Abundance dataset zero-adjusted in the same way. NMDS ordinations were also qualitatively

interpreted to assess whether the key aspects of the patterns of community change of individual rehabilitated sites were maintained across analyses, examining the relationships between individual points. This involved examining the ordinations of each simplified analysis and for each rehabilitated site interpreting the net direction of community change (converging, diverging or no net movement) and any notable characteristics of the pattern, such as apparent rate of change, changes in direction of community change over the course of the sampling period or general shape of the pattern. This method was employed to take into account features of convergence patterns which are not easily quantified but are nonetheless notable for interpretation purposes, particularly the relative positions of samples from the same site, which are not considered in the quantitative analyses employed. Procrustes analysis of NMDS ordinations examines the extent to which the positions of each point in the ordination differs between ordinations, rather than how specific points move relative to other specific points, and so cannot be used to assess the extent to which specific patterns of community change for individual rehabilitated sites are preserved in the simplified analyses, and thus the extent that they can be interpreted in the same way as ordination of Species Abundance data. This method was also employed in order to. Interpretations were done by a single author in order to ensure standards of interpretation were consistent across simplified analyses.

ARC-Dissimilarity results were compared at two levels, assessing fidelity of the simplified analyses' ability to reproduce rehabilitated site trend direction (convergent/divergent) and time to convergence. First, the estimated years to convergence of each rehabilitated site with each of the three reference sites at their respective mines was calculated from linear models. These times to convergence and from divergence are proxies, rather than true estimates – in reality we not expect rehabilitated site communities to become identical to the ARC but instead to level off at an earlier stage (as discussed in Chapter 2 Methods, Section 3). However, there is insufficient data to track the idealised rehabilitation scenario to the asymptote of the curve, so instead the of the unknown asymptote dissimilarity, the ARC is used as the end-point of recovery.

This allows for calculation of "years to convergence" as an intuitive value that incorporates not only the rate of convergence but also how dissimilar to the ARC a rehabilitated site was at the beginning of sampling, and so provides a proxy for the true recovery time that is comparable between sites. It also allows for equivalent values to be calculated for divergent sites – the initial dissimilarity is unknown, but using this approach the ARC can be used as an arbitrary starting point and their "years since divergence" (the inverse of years to convergence) can be calculated and compared to other estimated "years to convergence/from divergence". These estimates were calculated using formula:

$$\frac{site\ intercept}{-(site\ slope)}$$

Sites with negative intercepts were corrected to an intercept of 0.01 (1% dissimilarity), as dissimilarity values below 0 are not biologically possible. To assess reproduction of site trend direction, these 'years to convergence' results were then classified into convergent or divergent with the ARC and an unweighted Cohen's Kappa analysis was conducted to assess agreement between convergence/divergence classifications for each simplified analysis and Species Abundance analysis. Kappa scores were classified according to reliability based on the suggested thresholds provided in Landis & Koch (1977) (table 4). The number of matching classifications was also tallied to produce a raw percentage matching rate for each simplified analysis. Analysis was done in R v.3.4.3 using package *irr*.

Kappa score range	Agreement (Landis & Koch, 1977)	Reliability
<0	No agreement	No agreement
0-0.20	Slight Agreement	Extremely Poor
0.21-0.40	Fair Agreement	Poor
0.41-0.60	Moderate Agreement	Moderate
0.61-0.80	Substantial Agreement	Substantial
0.81-1.00	Almost perfect	Excellent

Table 4: *Classification of level of reliability of simplified analyses for replicating convergence/divergence results according to Cohen's Kappa scores, based on Landis & Koch, 1977.*

To assess the fidelity of the simplified analyses at the more precise level of estimating the years to convergence for each rehabilitated site, an intraclass correlation coefficient (ICC) analysis was used to assess the extent of agreement between each simplified analysis and Species Abundance analysis. Estimated Years to Convergence values were first transformed to their reciprocals (1/x). This was done in order to correct for the issue that, in the ARC-Dissimilarity model, the closer a site's slope is to horizontal (and therefore the slower the rate of change), the larger the "estimated years to/from convergence" values become, as the slower the rate of dissimilarity change, the longer it would take for sites to converge or the longer it would have been since diverging sites "diverged" from the reference site ARCs. This means that the greater the number of years to convergence or since divergence of a site, the smaller the difference between that value and its equivalent, so 3000 years to convergence is actually more similar to 3000 since divergence than 50 years to convergence is to 50 years since divergence. This creates problems when comparing convergence time estimates between Species Abundance and various simplified analyses, as there is sometimes disagreement between the two as to whether sites are converging or diverging, and though in reality the difference between larger values is smaller than the difference between smaller values, this inverted data structure cannot be properly analysed as statistical packages do not support it.

The simplest solution to this problem is to convert results to their reciprocals, resulting in a conventional data structure which can be properly analysed. This transformation is visualised in Appendix 2 (figs. 41 & 42). Reciprocal data was then analysed using a two-way mixed single-rater Intraclass Correlation Coefficient analysis of absolute agreement. ICC scores were classified by reliability based on the classification suggested in Koo & Li (2016) (table 5). Analysis was conducted in SPSS build 1.0.0.1126.

ICC Score	**Reliability**
<0.5	Poor
0.5-0.74	Moderate
0.75-0.89	Good
0.9-1.0	Excellent

Table 5: *Classification of level of reliability of simplified analyses for replicating "Estimated Years to Convergence" (EYCA) results, based on ICC scores. Scale taken from Koo & Li (2016).*

Results

1 - Callide Mine

1.1 – Overview

Procrustes analysis of NMDS ordinations (table 6) shows that Species Presence/Absence analysis (fig. 8) produced the most similar NMDS ordination to that displayed by Species Abundance analysis, while Large-Bodied Species Abundance analysis (fig. 7) produced the least similar ordination. Genus Abundance analysis (fig. 5) also produced a relatively similar ordination, while Functional Group Abundance analysis (fig. 6) sits in the middle of the pack.

Procrustes NMDS Comparisons	**Genus Abundance**	**Functional Group Abundances**	**Large-Bodied Species Abundance**	**Species Presence/ Absence**
Callide	9.491	12.34	19.2	5.493

Table 6: *Procrustes analysis results for comparison of simplified analyses to NMDS ordination of ant species abundance at Callide Mine. Lower Procrustes scores = greater similarity of ordination based on that approach to that produced by species abundance approach. Large-Bodied Species Abundance Species NMDS ordination is zero-adjusted and compared with a zero-adjusted Species Abundance ordination.*

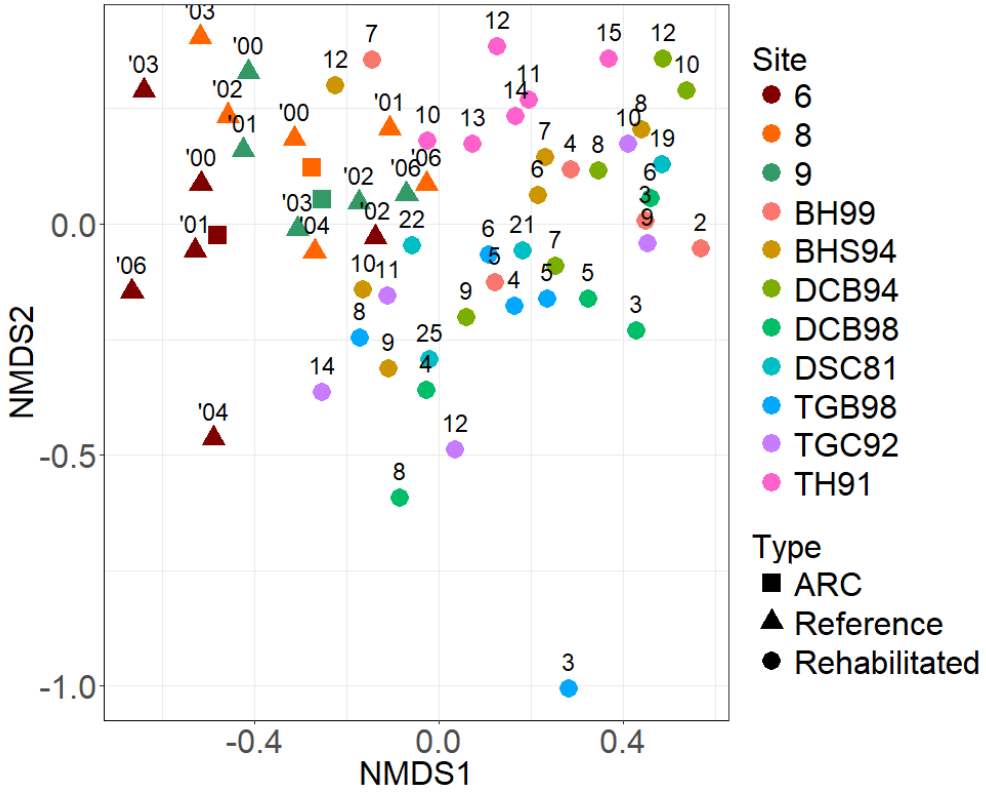


Figure 5: *Non-metric multidimensional scaling ordination of Callide Mine Sites based on ant genus abundances, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (2000-2006) for reference site samples. Two-dimensional stress = 0.172341*

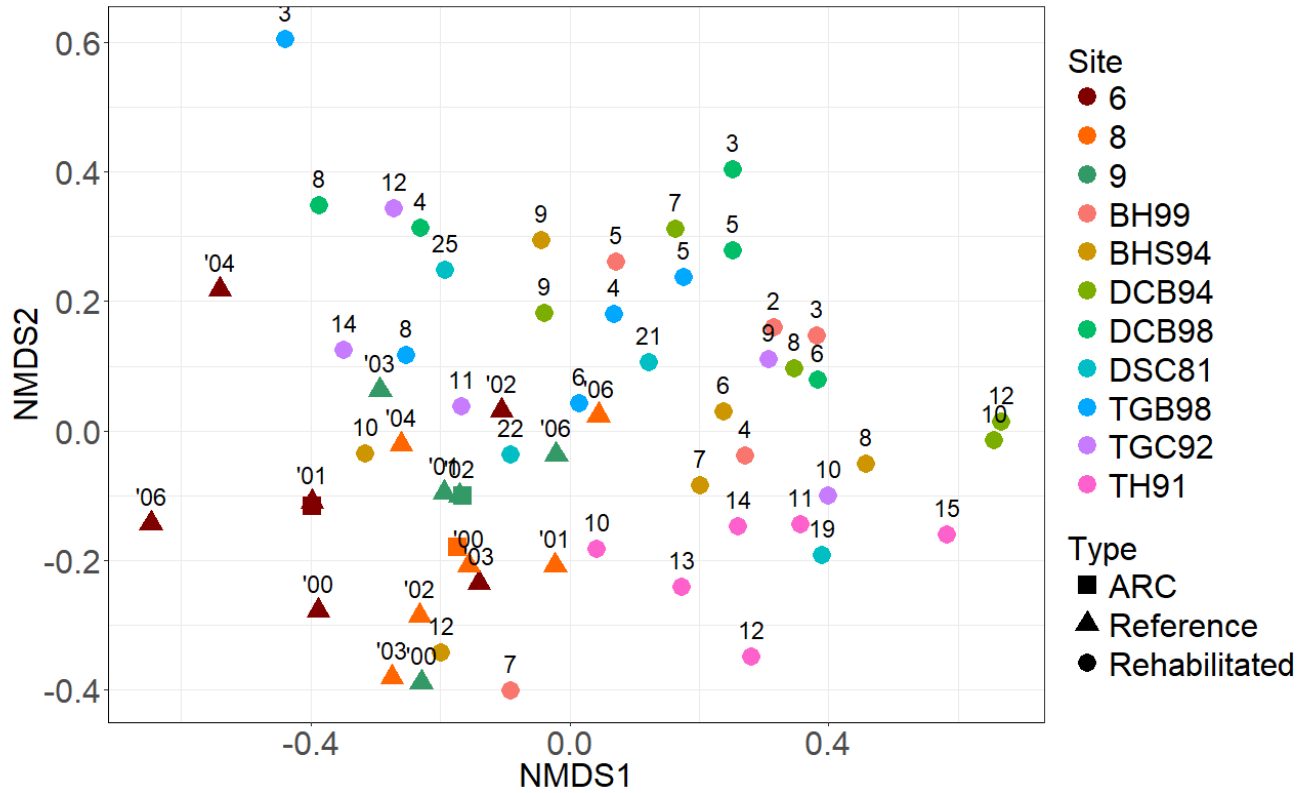


Figure 6: *Non-metric multidimensional scaling ordination of Callide Mine Sites based on ant functional group abundances, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (2000-2006) for reference site samples. Two-dimensional stress = 0.1625742*

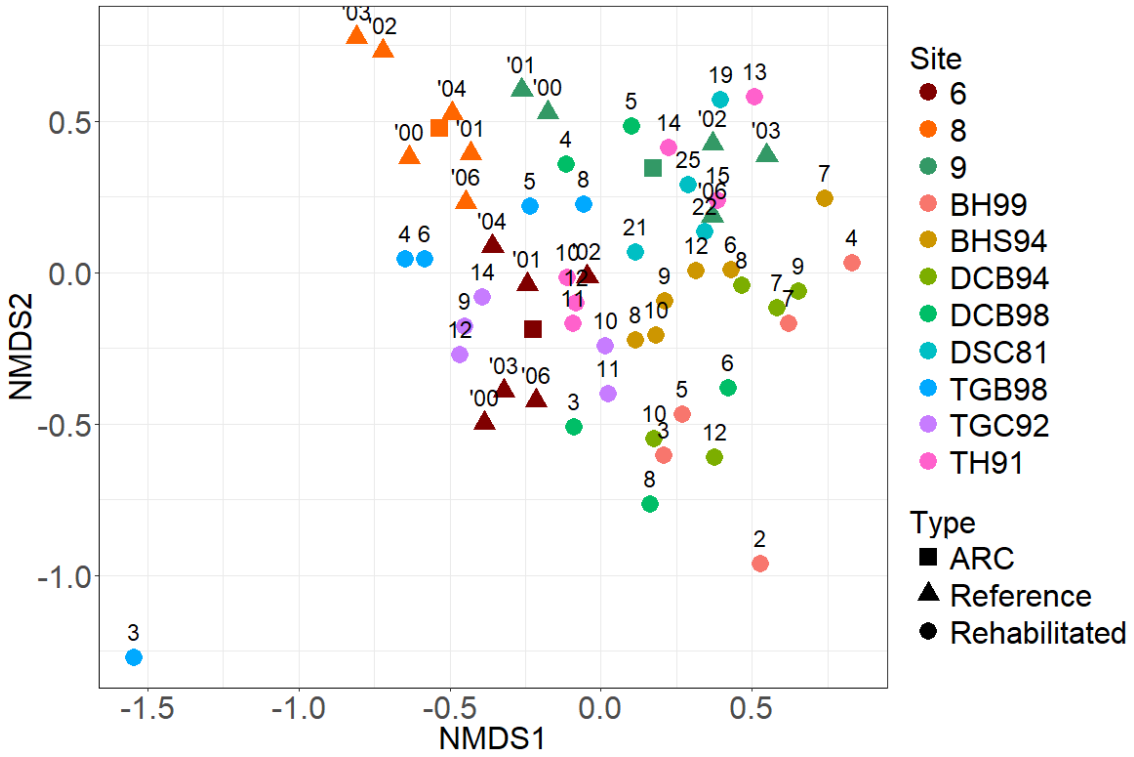


Figure 7: *Non-metric multidimensional scaling ordination of Callide Mine Sites based on large bodied ant species abundances, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (2000-2006) for reference site samples. Bray-Curtis pairwise site dissimilarities are zero-adjusted. Two-dimensional stress = 0.2203731*

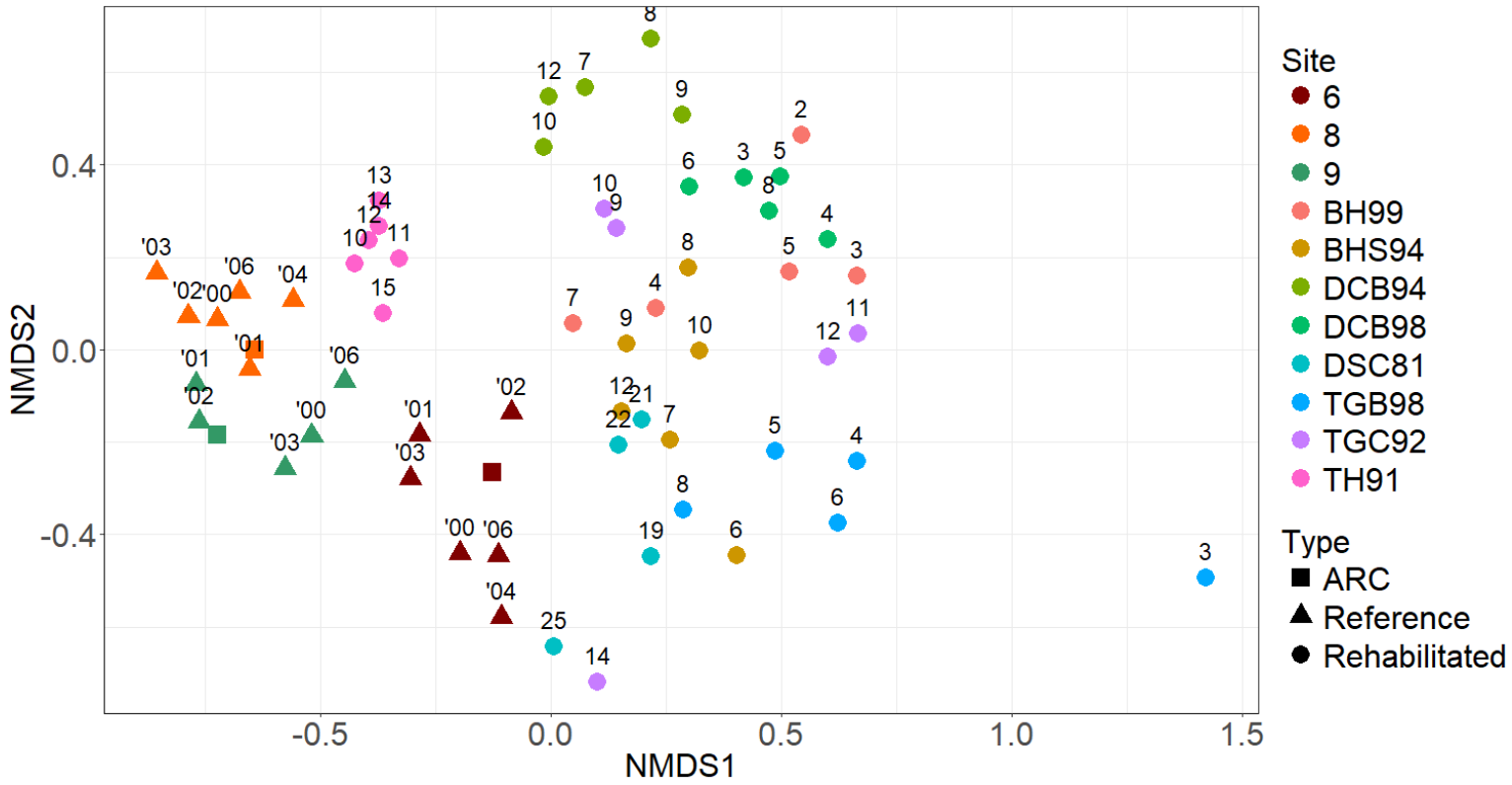


Figure 8: *Non-metric multidimensional scaling ordination of Callide Mine Sites based on ant species presence & absence, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (2000-2006) for reference site samples. Two-dimensional stress = 0.219685*

In contrast, comparison of the four different simplified analyses with Species Abundance analysis under the ARC-Dissimilarity approach shows that, at the level of classifying rehabilitated sites as either converging or diverging from the reference sites (table 7), all four sites had a relatively similar degree of agreement with Species Abundance analysis, with all four having levels of agreement significantly higher than chance. Three of the four analyses had Kappa scores falling into the "moderate reliability" category, while Species Presence/Absence fell just below the cut-off. However, like the other three simplified analyses it had high raw percentage matching of results for each site, over 80%. The rank order of the four analyses also differed, with Species Presence/Absence analysis having the lowest reliability of the four and Genus Abundance analysis also scoring relatively low, while Functional Group Abundance analysis was the most reliable. However, in practice there is little difference in Kappa reliability scores or percentage matching between the four simplified analyses.

Convergence/ Divergence	**Genus Abundance**	**Functional Group Abundance**	**Large-Bodied Species Abundance**	**Species Presence/ Absence**
Kappa	0.6**	0.69***	0.625**	0.571**
% Matching	83.33	87.5	83.33	83.33

Table 7: *Cohen's Kappa values of the four simplified analyses for matching site classification as converging or diverging from the ARC with results of Species Abundance analysis at Callide Mine. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001).*

At the level of estimating the years until community composition convergences with the ARC of each reference site (EYCA, table 3), results were substantially poorer, with only Species Presence/Absence analysis achieving significant

agreement with Species Abundance analysis. Despite this general poor performance, the lone successful simplified analysis, Species Presence/Analysis, had a high reliability score. The rankings of the four simplified analyses are in line with those of Procrustes analysis, although the relative agreement of each analysis does not directly correspond to their Procrustes scores.

EYCA	Genus Abundance		Functional Group Abundances		Large-Bodied Species Abundance		Species Presence/ Absence	
	ICC	Confidence Interval	ICC	Confidence Interval	ICC	Confidence Interval	ICC	Confidence Interval
Scores	0.288	-0.117-0.612	0.027	-0.393-0.426	0.012	-0.385-0.405	0.765***	0.525-0.891

Table 8: *Intraclass Correlation Coefficients (ICC) of the four simplified analyses at Callide Mine for matching the estimated years to convergence with ARCs (EYCA) of each reference site with results of Species Abundance analysis at Callide Mine. Colours indicate reliability (Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = ≤0.01, *** = ≤0.001).*

The three analyses, Procrustes, Convergence/Divergence Kappa and EYCA ICC, present quite different pictures of which of the four simplified analyses may serve as surrogates for Species Abundance analysis. Based on Procrustes scores, Species Presence/Absence would appear to show close concordance with the NMDS ordination produced by Species Abundance analysis, and is the strongest analysis by far for EYCA analysis, though still only at moderate levels of agreement. but in convergence/divergence models it is the poorest-performing at replicating the results of Species Abundance analysis and has the lowest agreement of the four analyses, though only by a small margin. In contrast, Functional Group Abundance analysis is the best-performing simplified analysis under ARC-Dissimilarity at both levels, but is 3rd-most similar in Procrustes analysis with a score that places it in the middle of the score range of scores and performs poorly at the EYCA level. Large-Bodied Species Abundance analysis shows a relatively idiosyncratic set of scores, performing worst in Procrustes and EYCA analyses, but having the second-highest agreement at the convergence/divergence level of ARC-Dissimilarity analysis. Performance of the four analyses at the convergence/divergence level is relatively high – despite only being poorly or moderately more reliable than random classification, for all four analyses this corresponds to raw % agreement in excess of 80%, so all four analyses could be viable as surrogate analyses. But at the EYCA level it is a different story, with only Species Presence/Absence having significant agreement reliability with Species Abundance analysis.

1.2 – Genus Abundance

Genus Abundance analysis remains in the middle ground in terms of relative agreement with Species Abundance analysis, being neither the best or worst simplified analysis in any of the three analyses used to assess agreement (table 9). However, in ARC-Dissimilarity its agreement reliability with Species Abundance analysis is only significant at the convergence/divergence level ($z(23,23) = 3$, $p = 2.7x10^{-3}$), at the EYCA level it does not achieve significance ($F(23,23) = 1.81$, $p = 0.081$).

	Procrustes	**Convergence/ Divergence**		**EYCA**	
Scores	9.491	0.6**	83.33%	0.288	-0.117-0.612
Rank	2	3		2	

Table 9: *Summary of scores and ranks of Genus Abundance analysis in each of the three agreement analyses. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001).*

An examination of the ARC-Dissimilarity convergence patterns themselves reveals a much greater degree of homogeneity between the patterns relative to the three reference sites (fig. 9) than those of Species Abundance analysis, and the dissimilarities between rehabilitated sites and the ARC are generally lower than in Species Abundance analysis (fig. 2 (Chapter 2 Results)). This is also apparent in the ordination for Genus Abundance analysis (fig. 5), in which the spread of site samples is much more condensed and individual points lie closer together in ordination than in the ordination for Species Abundance analysis (fig. 1 (Chapter 2 Results)), indicating reduced dissimilarity between them. The three reference sites in particular occur much more closely together, with greater overlap, than in the ordination of Species Abundance analysis.

Figure 9 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 6 (9a), 8 (9b) and 9 (9c) with site age since rehabilitation (years) at Callide Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 9a: *Convergence with Site 6, Callide Mine*

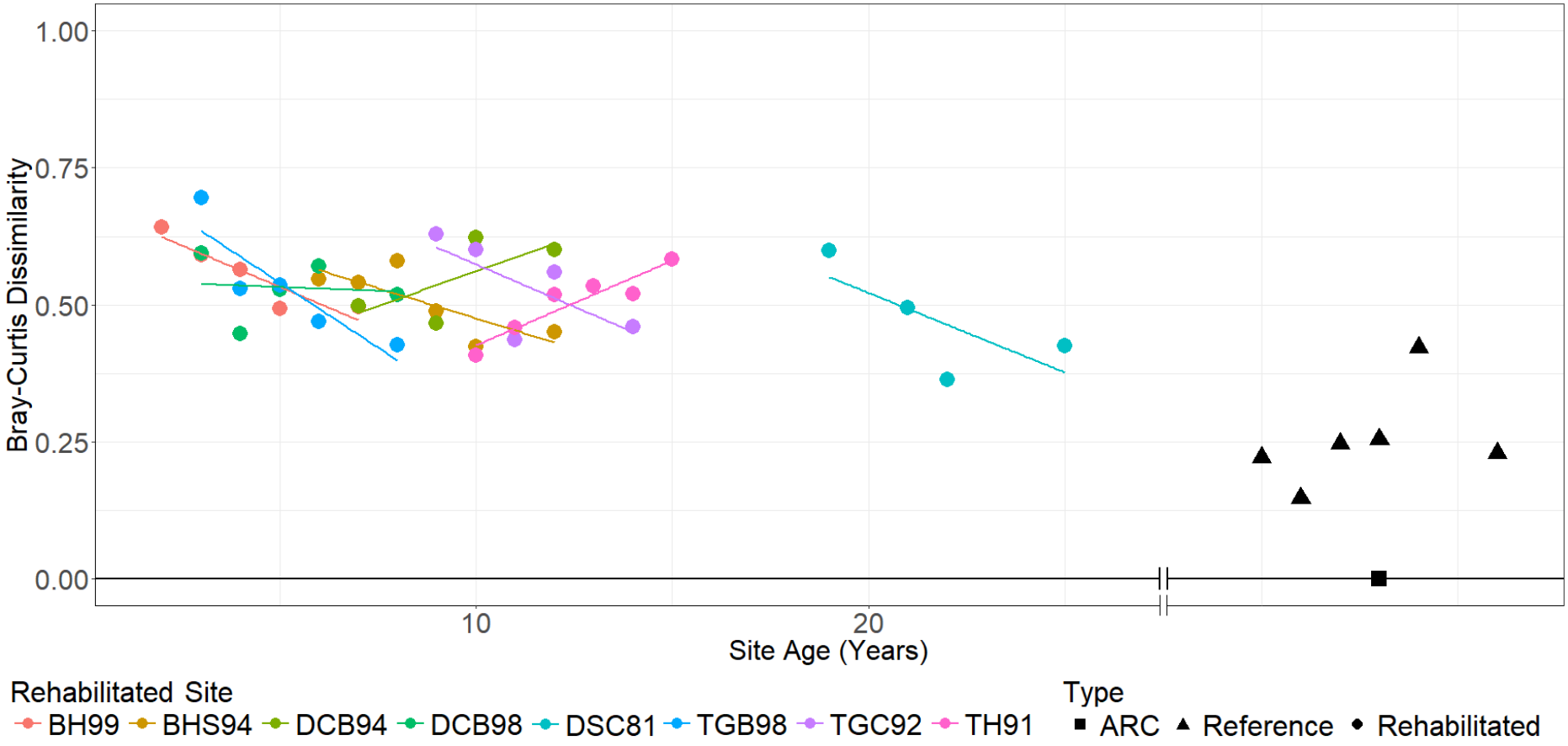


Fig 9b: *Convergence with Site 8, Callide Mine*

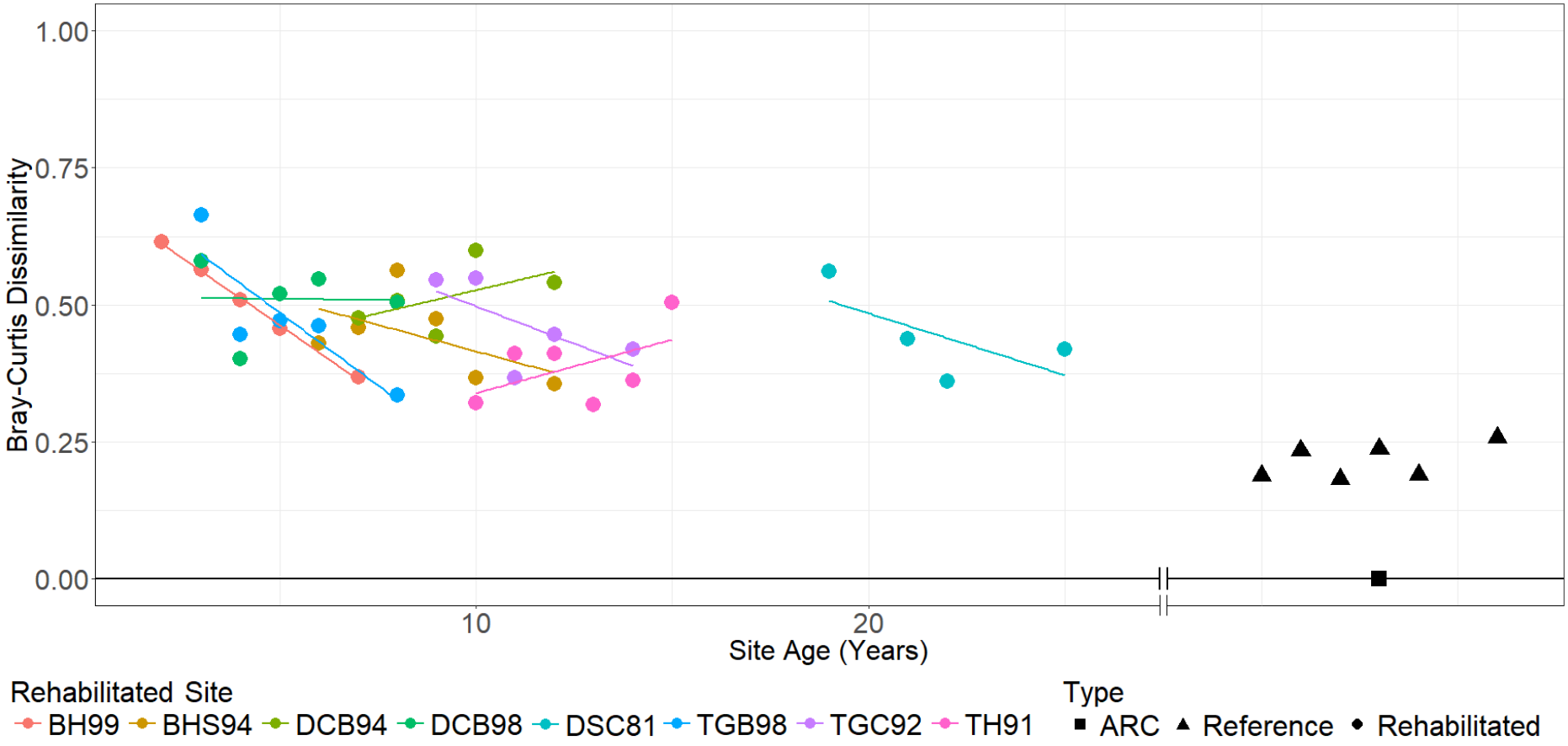
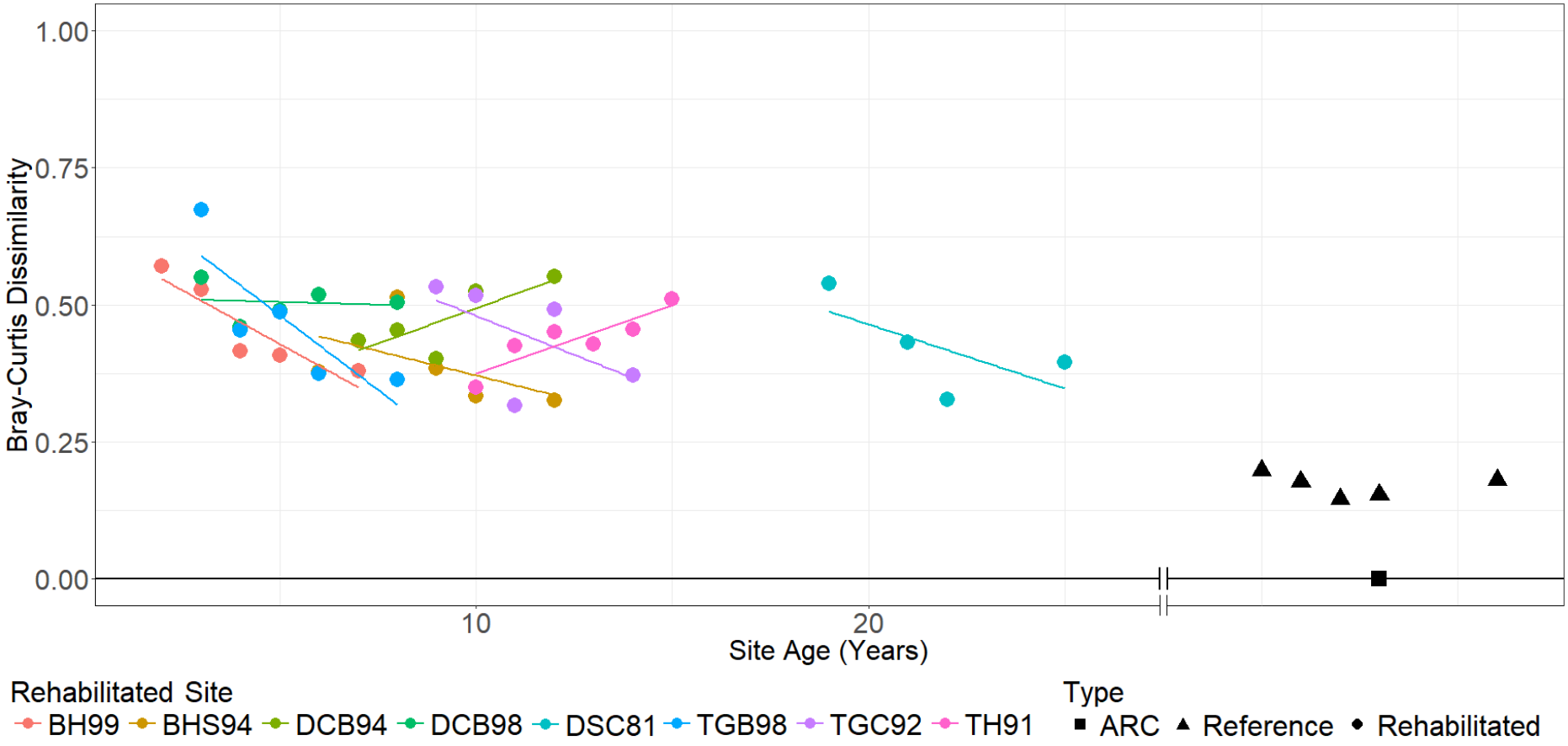


Fig 9c: *Convergence with Site 9, Callide Mine*



1.3 – Functional Group Abundance

Functional Group Abundance analysis had the highest agreement with Species Abundance Analysis based on ARC-Dissimilarity data for convergence/divergence ($z(23,23) = 3.55$, $p = 3.79 \times 10^{-4}$), though only by a relatively small margin, but was non-significant for EYCA ($F(23,23) = 1.053$, $p = 0.452$) (table 10). In Procrustes results, it was third with a score placing almost exactly midway between the lowest and highest scores (range = 5.493 – 19.2, mid-point = 12.3465).

	Procrustes	**Convergence/ Divergence**		**EYCA**	
Scores	12.34	0.69***	87.5%	0.027	-0.393-0.426
Rank	3	1		3	

Table 10: *Summary of scores and ranks of Functional Group Abundance analysis at Callide Mine in each of the three agreement analyses. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001).*

The ARC-Dissimilarity convergence patterns of Functional Group Abundance analysis (fig. 10) show a similar pattern to that of Genus Abundance analysis (fig. 9), being relatively homogeneous and resembling the Species Abundance analysis convergence patterns of reference site 6 The overall dissimilarity of rehabilitated sites to reference sites is even lower for functional groups than for genera however, such that many of the rehabilitated sites overlap with the reference site samples in ARC-Dissimilarity by the end of the sampling period, in marked contrast with Species Abundance analysis, where the only overlap at Callide Mine was of DSC81 in sampling year 3 with reference site 6, and TGC92 comes very close but does not ultimately fall within the range of dissimilarities displayed by the reference sites.

The NMDS ordinations for Genus (fig. 5) and Functional Group Abundance (fig. 6) analyses are likewise very similar in the relative positioning of site samples, reflecting the similarities of their ARC-Dissimilarity results (figs. 9 & 10), in particular the greatly reduced dissimilarity between the three reference sites and the resultant homogeneity of ARC-Dissimilarity convergence patterns on each of the three reference sites.

Figure 10 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 6 (10a), 8 (10b) and 9 (10c) with site age since rehabilitation (years) at Callide Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 10a: *Convergence with Site 6, Callide Mine*

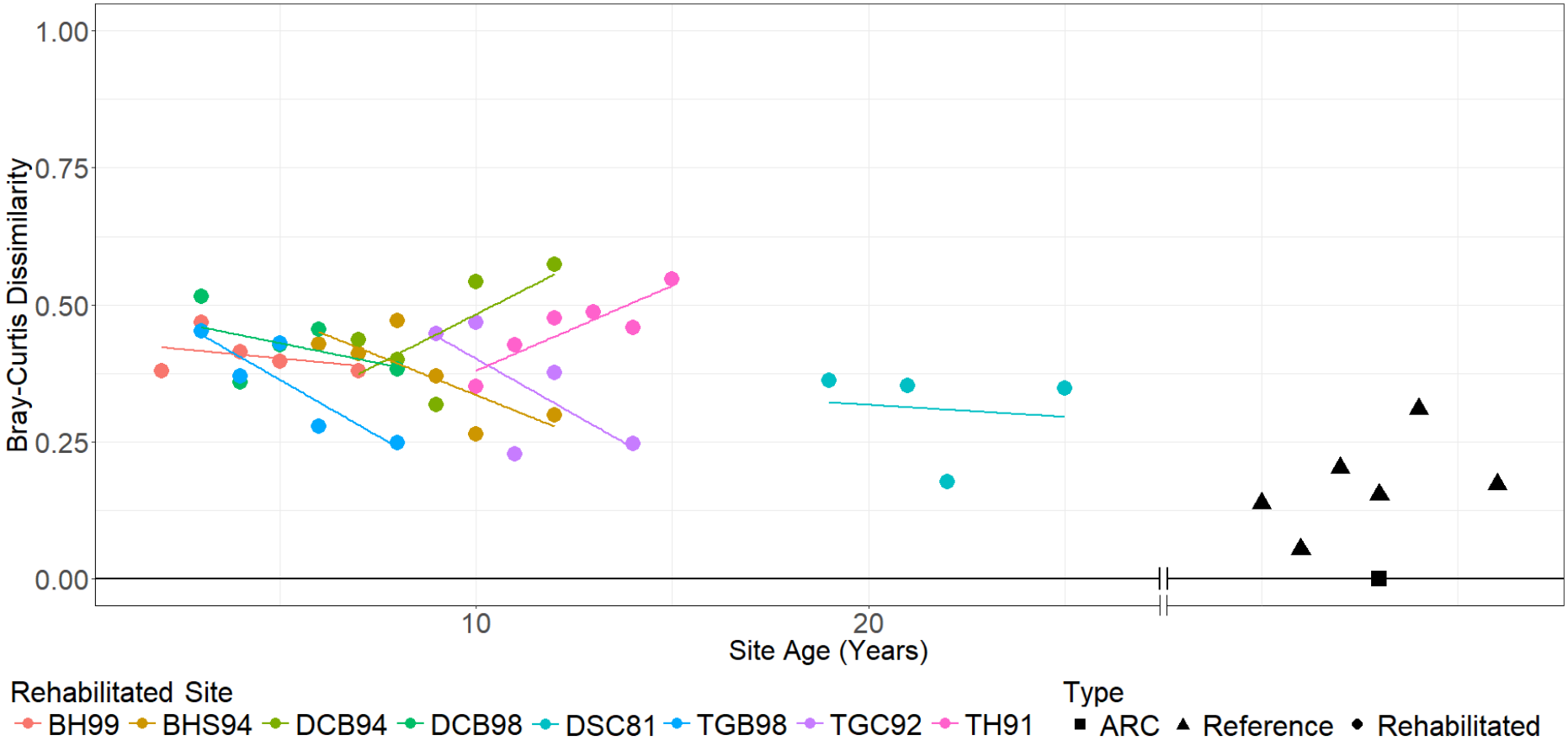


Fig 10b: *Convergence with Site 8, Callide Mine*

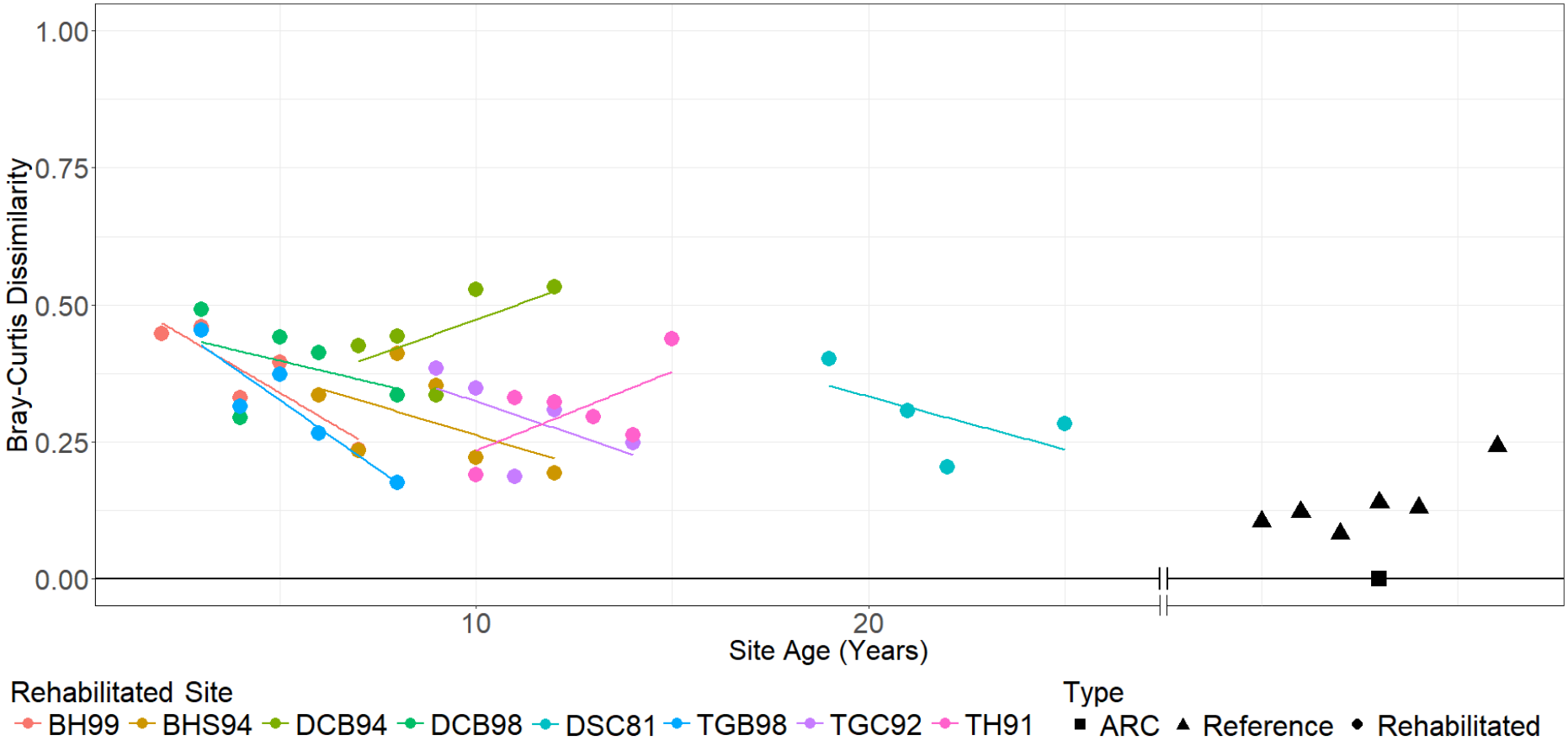
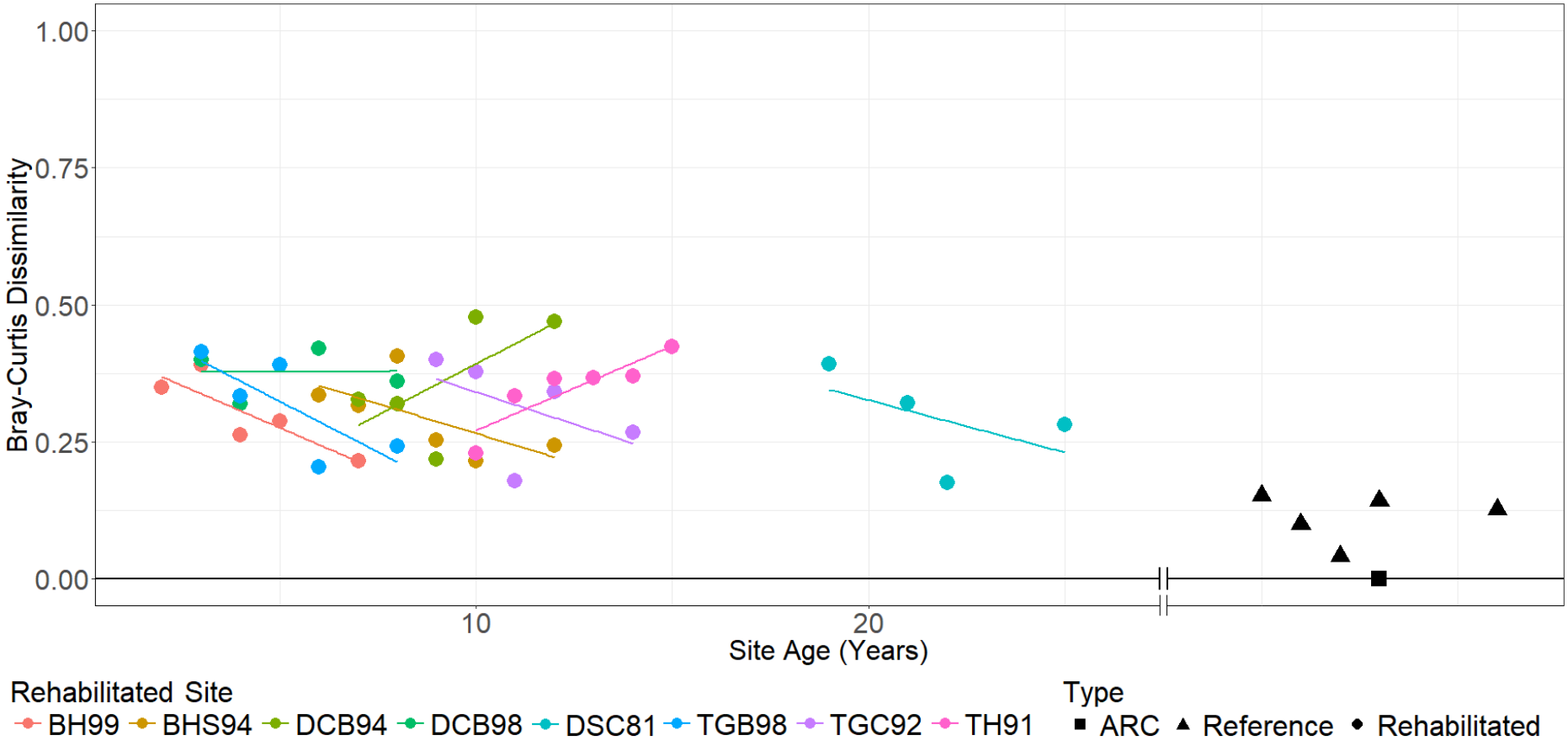


Fig 10c: *Convergence with Site 9, Callide Mine*



Section 1.4 – Large-Bodied Species Abundance

Large-Bodied Species Abundance analysis has a varied success rate across the three analyses (table 11), having the highest Procrustes score and a non-significant ICC score in EYCA analysis ($F(23,23) = 1.024$, $p = 0.478$), but achieving moderate rates of agreement with Species Abundance analysis on convergence/divergence similar to those of the other three analyses ($z(23,23) = 3.06$, $p = 2.2x10^{-2}$).

	Procrustes	**Convergence/ Divergence**		**EYCA**	
Scores	19.2	0.625**	83.33%	0.012	-0.385-0.405
Rank	4	2		4	

Table 11: *Summary of scores and ranks of Large-Bodied Species Abundance analysis at Callide Mine in each of the three agreement analyses. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001).*

Unlike the ARC-Dissimilarity results of Genus Abundance and Functional Group Abundance analyses, ARC-Dissimilarity results for each rehabilitated site are not homogenised across reference sites, either in terms of convergence patterns or overall dissimilarity of rehabilitated sites to reference sites, and do not bear a particularly close resemblance to any one reference site's set of convergence patterns in Species Abundance analysis (fig. 11).

Figure 11 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 6 (11a), 8 (11b) and 9 (11c) with site age since rehabilitation (years) at Callide Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 11a: *Convergence with Site 6, Callide Mine*

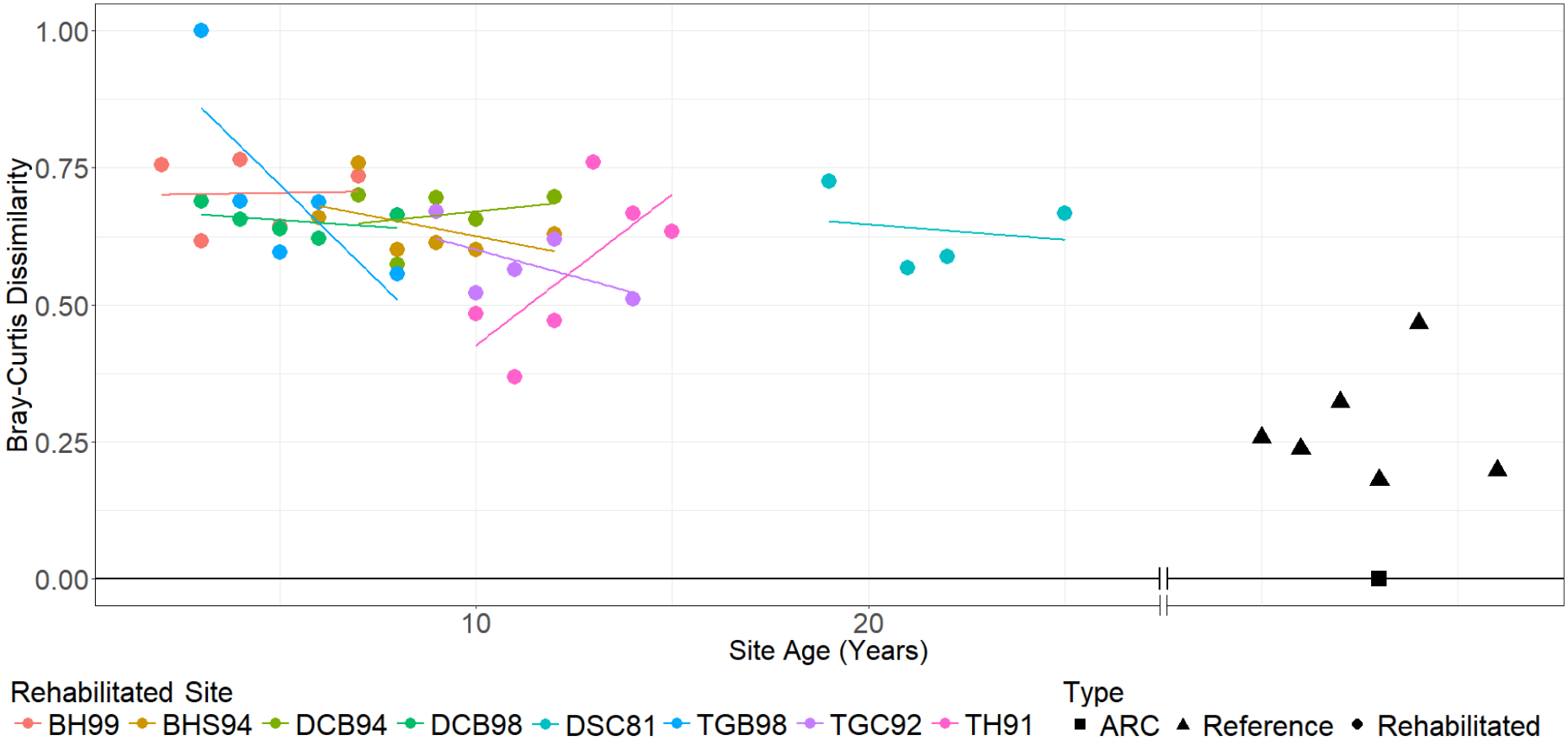


Fig 11b: *Convergence with Site 8, Callide Mine*

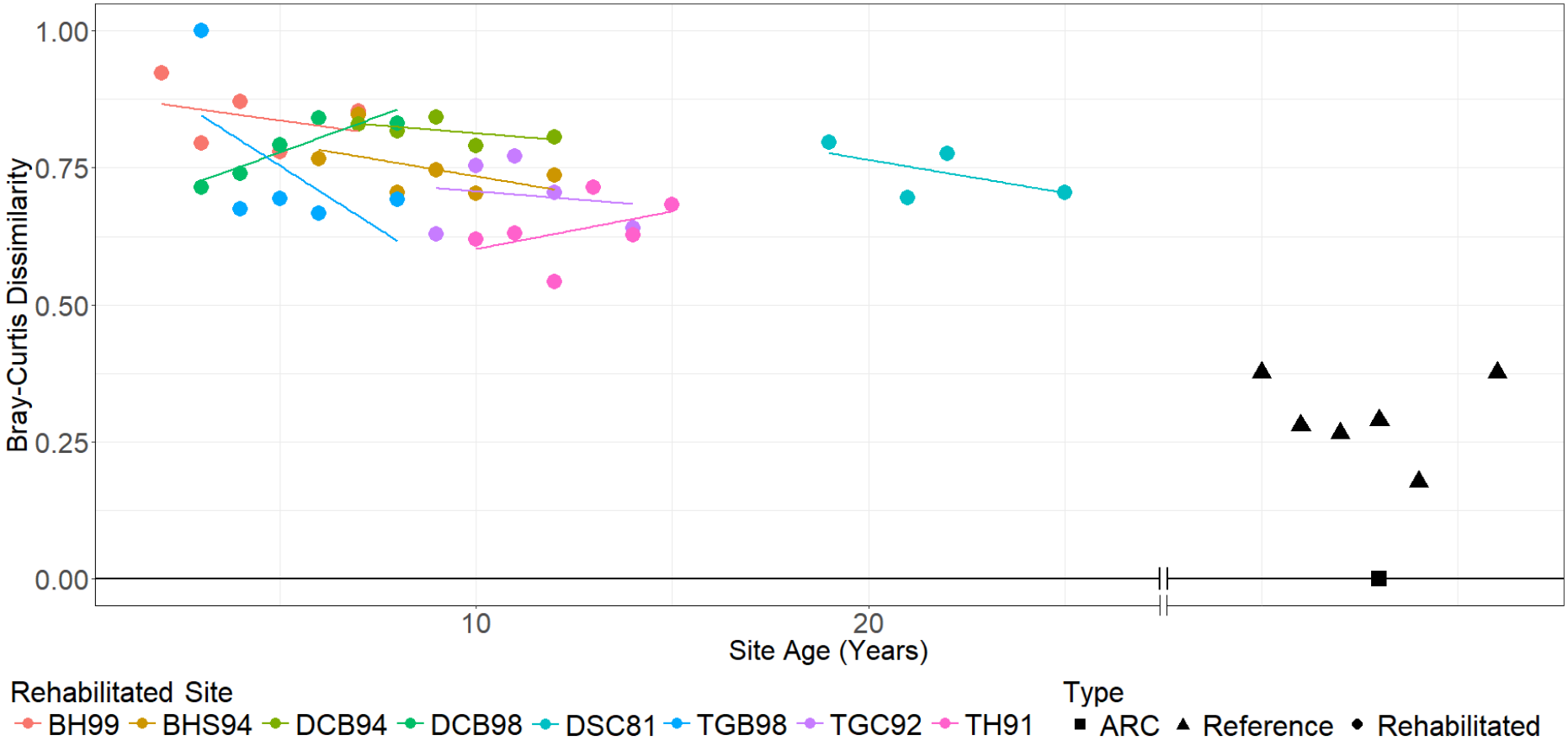
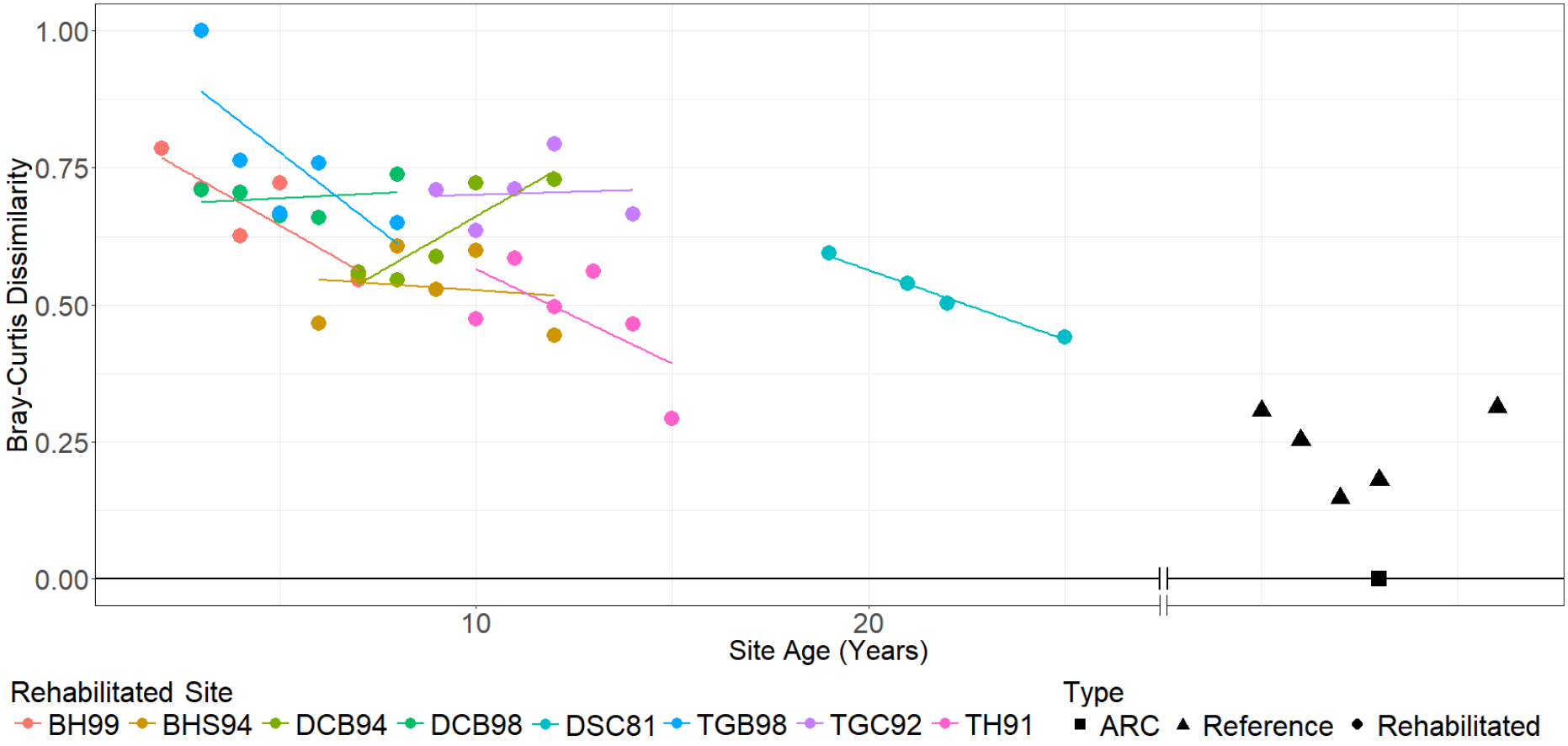


Fig 11c: *Convergence with Site 9, Callide Mine*



Looking at the NMDS ordination for Large-Bodied Species Abundance species (fig. 7), we see a very different ordination to that of Species Abundance (fig. 1 (Chapter 2 Results)), corresponding to its having the highest Procrustes sum-of-squares of the four analyses (table 11). In this ordination reference and rehabilitated sites overlap substantially, with reference sites 6 & 9 overlapping with the rehabilitated sites almost completely. However, the lack of overlap of reference and rehabilitated sites in ARC-Dissimilarity (fig. 11) suggests that in this ordination, the primary axis of convergence is not the focus of this 2D plot, rather than being a result of rehabilitated sites and reference sites being the same in terms of Large-Bodied Species Abundance Species abundance. Instead, this ordination pattern is predominantly a result of the positioning of TGB98 at 3 years post-rehabilitation. This sample lacks any Large-Bodied Species Abundance species, meaning it is the most dissimilar sample to all other samples at Callide Mine on average by a large margin, and subsequently the dissimilarity between TGB98 at 3 years post-rehabilitation and all other samples displaces the dissimilarity between reference and rehabilitated sites, the primary axis of convergence, as the main axis of dissimilarity around which the ordination is structured, as discussed in chapter 2 (discussion, section 2). It is very difficult to interpret this ordination in terms of convergence/divergence given the primary axis of convergence is not one of the two axes of greatest variance (NMDS 1 & 2), or to relate it in any concrete way to the ARC-Dissimilarity trends. Movement of rehabilitated sites appears to be largely aligned to a reference site 6 – reference site 9 axis (bottom left of the cluster (reference 6) to top right (reference 9), approximately along line X=Y) with little convergent or divergent movement relative to the reference sites as a whole, but movement from one side of the cluster to the other. However, the only sites this movement translates into movement away from one reference site towards the other for in ARC-Dissimilarity are BH99 and TH91, suggesting this does not reflect the true extent of convergence or divergence for most sites.

1.5 – Species Presence/Absence

Species Presence/Absence had the lowest Procrustes score of the four analyses (table 12) and is the only simplified analysis to have a significant EYCA ICC score ($F(23,23) = 8.187$, $p = 1.92x10^{-6}$), or one with good reliability. In contrast, its convergence/divergence Kappa score is the lowest of the four simplified analyses, falling just below the "poor reliability" threshold (0.60) ($z(23,23) = 3.1$, $p = 1.95x10^{-3}$). However, this relatively poor Kappa score corresponds to a high percentage agreement of 83%, making Species Presence/Absence analysis a consistently strong simplified analysis at Callide Mine.

	Procrustes	**Convergence/ Divergence**		**EYCA**	
Scores	5.493	0.571**	83.33%	0.765***	0.525-0.891
Rank	1	4		1	

Table 12: *Summary of scores and ranks of Species Presence/Absence analysis at Callide Mine in each of the three agreement analyses. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001).*

ARC-Dissimilarity results for Species Presence/Absence (fig. 12) reflect these strong agreement scores. They are heterogenous and convergence patterns for each reference site show broad similarities to the convergence patterns of the respective sites under Species Abundance analysis in terms of general trends in convergence, positions of individual site-samples and overall dissimilarity of reference and rehabilitated sites, although they are not identical, particularly the convergence pattern for reference site 6, where TH91 is convergent when it is divergent under Species Abundance analysis, and the slope of TGC92's convergence is much shallower under Species Presence/Absence than Species Abundance.

The NMDS ordination (fig. 8) for Species Presence/Absence also shows a clear resemblance to the Species Abundance ordination (fig. 1 (Chapter 2 Results)), corresponding to the lowest Procrustes sum-of-squares being between these two simplified analyses. The reference sites are clearly not clustered together, as expected from the lack of homogeneity across results relative to the three reference sites. Reference and rehabilitated sites are more clearly separated and rehabilitated sites are slightly more clustered, making Species Presence/Abundance slightly more conservative in terms of gauging rehabilitated site restoration than Species Abundance. Patterns in the Species Presence/Absence ordination may vary substantially compared to the Species Abundance ordination, but the overall pattern of convergence/divergence for each site, with the exception of DCB94, appears to be unaffected, judging by the high reliability of agreement between sites.

Figure 12 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 6 (12a), 8 (12b) and 9 (12c) with site age since rehabilitation (years) at Callide Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 12a: *Convergence with Site 6, Callide Mine*

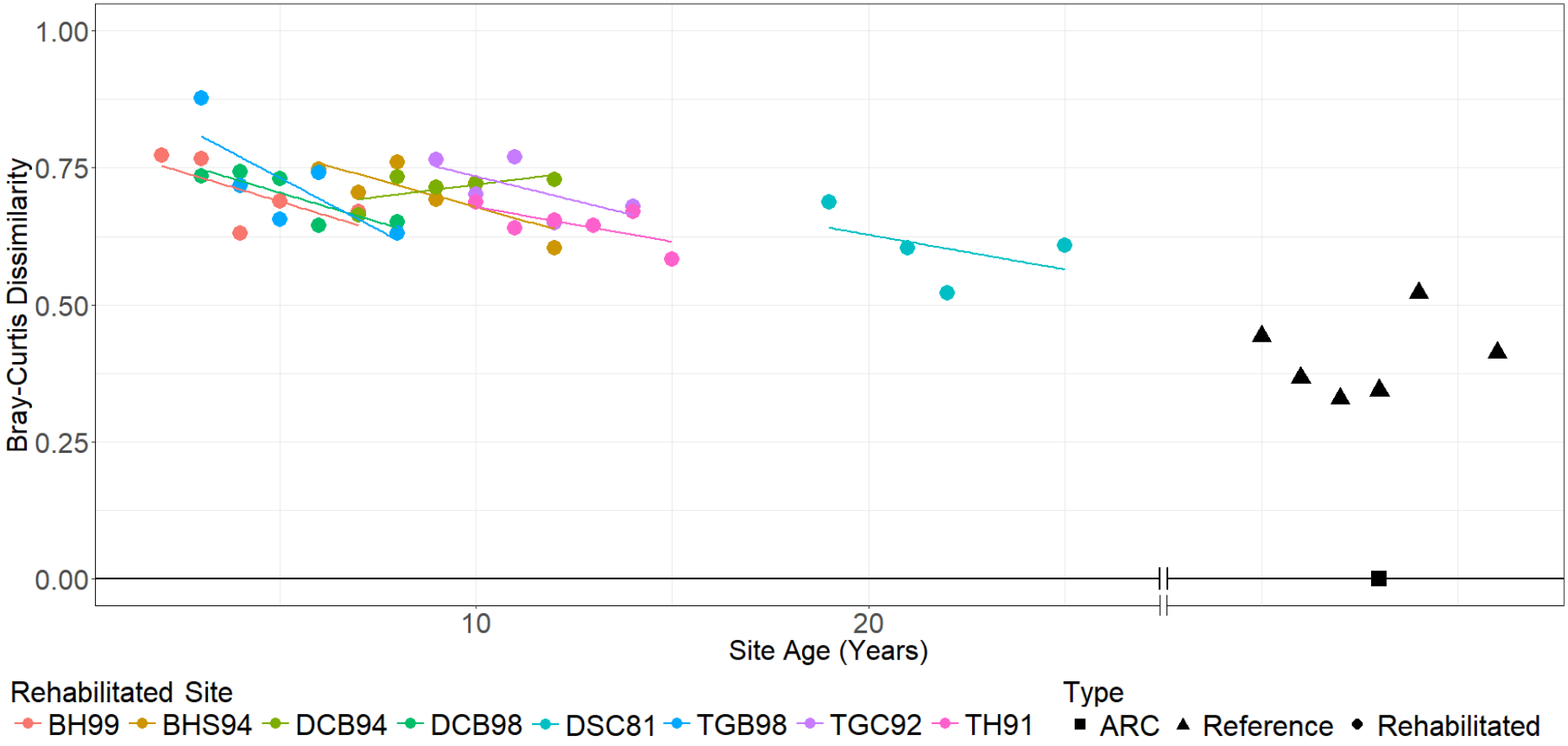


Fig 12b: *Convergence with Site 8, Callide Mine*

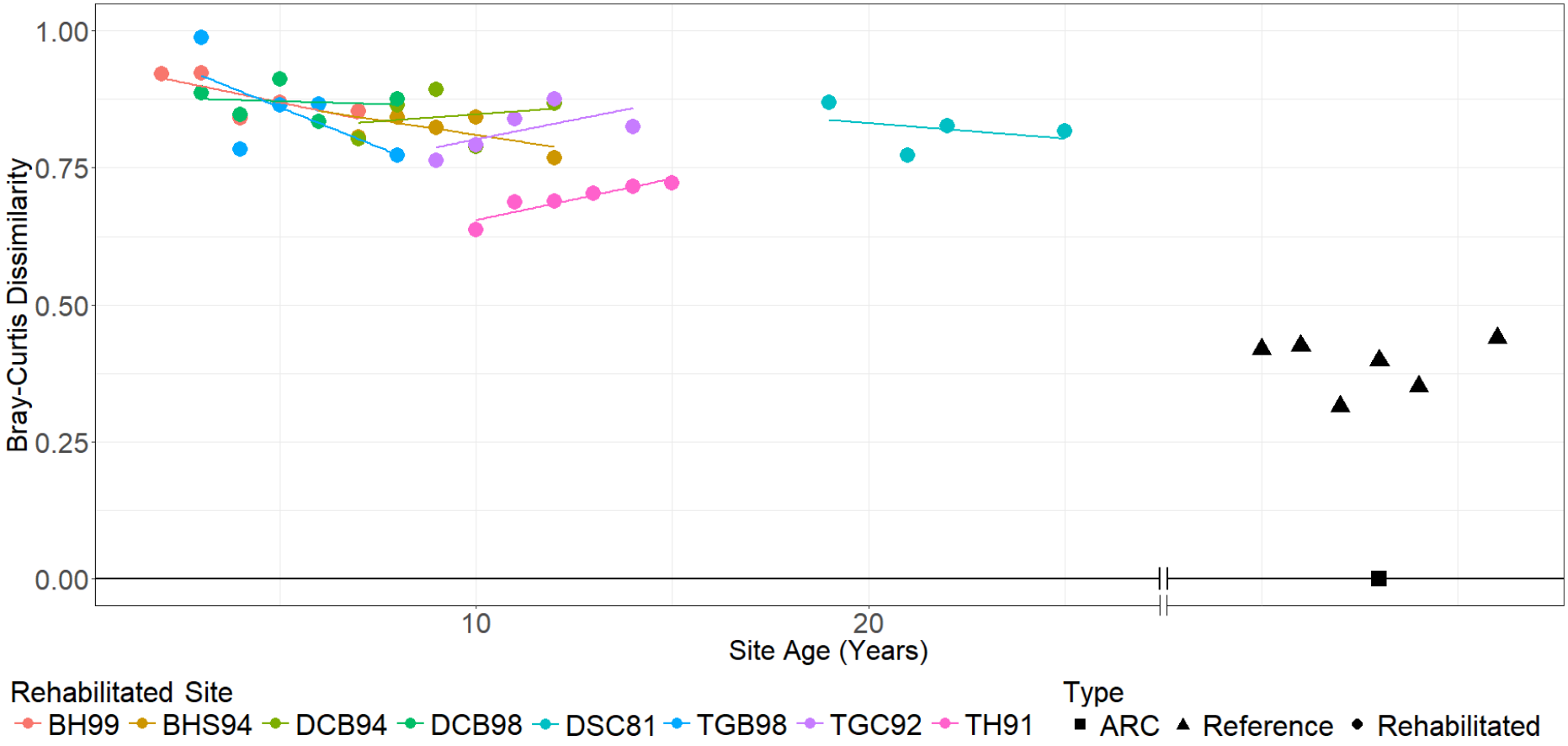
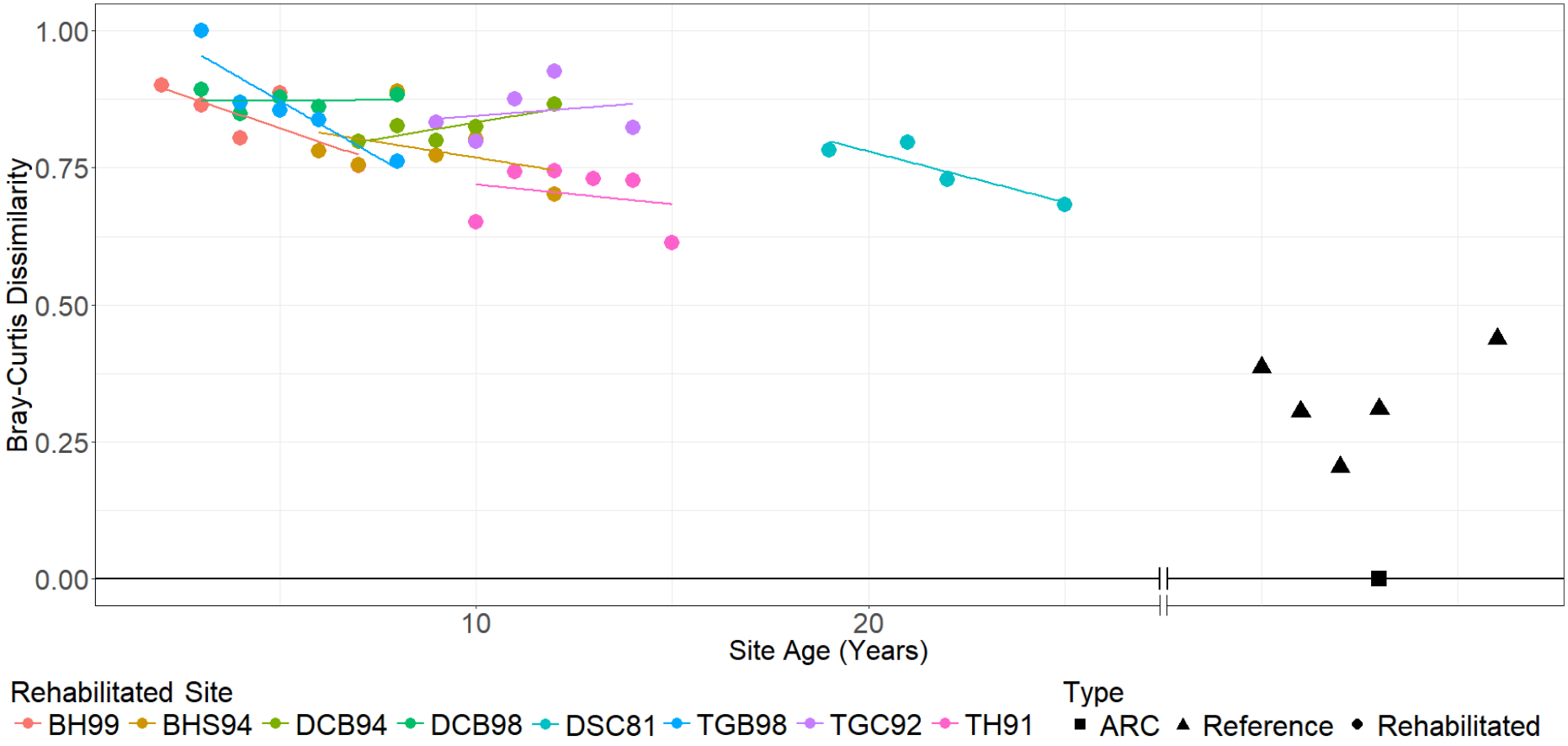


Fig 12c: *Convergence with Site 9, Callide Mine*



1.6 – Rehabilitated Site Trends (NMDS Ordination)

In qualitative interpretation of rehabilitated site trajectories and directions in NMDS Ordinations (table 13), Genus Abundance analysis was the most reliable analysis, producing similar directional results to Species Abundance for every rehabilitated site. Functional Group Abundance analysis was the second most reliable, producing the same results of Species Abundance analysis for 7/8 sites. Species Presence/Absence analysis was slightly less reliable, reproducing overall results for six of the eight sites but failing to replicate the clear divergence of DCB94 and only partially retaining the trend for DSC81, placing its starting position much closer to reference site 6 than in Species Abundance. Large-Bodied Species Abundance Species analysis was the least reliable, retaining directional trends only for TGB98 and DCB98, though the extent to which the relative positions of sites contribute to this low replicability is unclear when the major defining feature of this ordination is the position of the first sample taken from TGB98 and its obscuring of the primary axis of convergence, which is much better-preserved in ordinations of other three simplified analyses. More minor differences were also apparent in the ordinations of the four simplified analyses, which did not alter the patterns enough to change the general trend of convergence but did change the story. For instance, BH99 was portrayed in Genus Abundance and Functional Group Abundance analysis as overlapping with reference sites, and potentially even heading into divergence in the future (or, given the zig-zag trajectory previously, about to move deeper into convergence with reference sites). BHS94 also clearly overlapped with reference sites in ordination space in Genus Abundance and Functional Group Abundance analysis ordinations, which it came close to doing in Species Abundance but did not actually reach. Similarly, TH91 groups with reference sites across Species Abundance, Genus Abundance, Functional Group Abundance and Species Presence/Absence analyses, but is diverging much more rapidly in Genus Abundance and Functional Group Abundance analyses, whereas it is relatively static in Species Abundance and Species Presence/Absence analysis. More broadly, Species Presence/Absence analysis adheres to the overall picture of the Species Abundance ordination much more closely than Genus Abundance or

Functional Group Abundance analysis, the direction of DCB94 and DSC81 notwithstanding. Therefore, it is important to note that even when trends match up in the broad sense, there are variations in nuance which, without knowing what the Species Abundance pattern looks like, may cause problems for practitioners.

Table 13 (next page): *Qualitative interpretation of rehabilitated site community trends in terms of convergence/divergence with reference sites at Callide Mine, based on NMDS ordination for each analysis. Large-bodied Species analysis ordination is based on zero-adjusted Bray-Curtis dissimilarities.*

	Original Analysis	*Simplified analyses*			
Rehabilitated Sites	***Species Abundance (fig. 1 (Chapter 2 Results))***	***Genus Abundance (fig. 5)***	***Functional Group Abundance (fig. 6)***	***Large-Bodied Species Abundance Species Abundance (fig. 7)***	***Species Presence/Absence (fig. 8)***
BH99	**Steady convergence towards reference sites**	Steady convergence towards reference sites, overlaps with reference sites by 2006	Steady convergence towards reference sites; overlaps with reference sites by 2006	No convergence on reference sites as group after 2001-02, moves across sample cluster from reference site 6 side to reference site 9 side	Erratic convergence on reference sites, doubles back on self in 2003-4. Net movement is convergent
BSH94	**Initial erratic convergence towards reference sites, consistent from 2003 onwards.**	Initial erratic convergence towards reference sites, consistent from 2003 onwards, overlaps with reference sites in 2006	Initial erratic convergence towards reference sites, consistent from 2003 onwards, overlaps with reference sites by 2006	Erratic, no net convergence on reference sites	Erratic convergence on reference sites

DCB94	**Steady divergence from reference sites**	Steady divergence from reference sites	Steady divergence from reference sites	No convergence on reference sites, moves from reference site 9 side of sample cluster to reference 6 side.	No convergence on reference sites
DCB98	**No convergence on reference sites**	No convergence on reference sites	Erratic, slight convergence on reference sites	Circular movement that takes it close to all reference sites, net movement is divergent from reference sites 6 & 8, but neutral relative to site 9	No convergence on reference sites

DSC81	**On trajectory that intersects with reference site 6 in 2003-2006 but if continued would become divergent**	On trajectory that comes close to convergence with reference site 6 in 2003 but if continued would become divergent	On trajectory that intersects with reference site 6 in 2003 but continues on to divergence	Net convergence on reference sites, particularly reference site 9, with which it overlaps	Erratic, net movement comes close to intersecting with reference site 6 but if continued would become divergent
TGB98	**Rapid convergence in 2001-02, 2004-06**	Rapid convergence in 2001-02, little convergence since (except on reference site 6 in 2004-06)	Rapid convergence in 2001-02, 2003-04, converging on reference site 6 in 2004-06	Rapid convergence in 2001-02, erratic with no net movement relative to reference sites 6 or 8 but convergent on reference site 9	Rapid convergence in 2001-02, relatively consistent convergence on reference sites

TGC92	**Rapid movement on trajectory parallel to the reference sites, no convergence with reference sites on primary axis of convergence**	Rapid movement on trajectory parallel to the reference sites, no convergence with reference sites on primary axis of convergence	Rapid movement on trajectory parallel to the reference sites, no convergence with reference sites on primary axis of convergence	Erratic, no net movement, overlaps with reference site 6	Trajectory initially strongly divergent and then enters on trajectory that passes by close by reference site 6 but is ultimately divergent in other direction
TH91	**Groups with reference sites. Very slow divergence from reference sites**	Groups with reference sites. Steady divergence from reference sites	Groups with reference sites. Steady divergence from reference sites	Erratic, diverges from overlapping with reference site 6 to overlapping with reference site 9	Groups with reference sites, slow but steady divergence up until 2005-06, when it converges strongly so net movement is convergent

1.7. – Summary (Callide Mine)

Out of the four simplified analyses, only Species Presence/Absence reliably agreed with Species Abundance analysis at Callide Mine at all levels of analysis. The other three simplified analyses achieved similarly high rates of percentage agreement for ARC-Dissimilarity convergence/divergence, in excess of 80%, and had correspondingly similar Kappa scores. In NMDS Interpretation, Genus and Functional Group Abundance analyses actually outperformed Species Presence/Absence, despite their lower Procrustes scores. But only Species Presence/Absence scored highly in ICC analysis of estimated years to convergence, whereas the other three analyses had no correlation with the results of Species Abundance. So while all four simplified analyses are able to replicate broad trends in community change at the rehabilitated sites, there is a dropping off of simplified analyses as the required level of precision of trend replication increases, until at the level of highest precision studied, estimating time to convergence, only Species Presence/Absence has sufficiently precise fidelity to reliably serve as a surrogate measure for Species Abundance analysis.

Section 2 - German Creek Mine

2.1 – Overview

Procrustes analysis of NMDS ordinations (table 14) shows substantially lower resemblance between NMDS ordination of Species Abundance analysis (fig. 3 (Chapter 2 Results)) and the ordinations of the four simplified analyses than at Callide Mine, with the Procrustes results for Genus (fig. 13), Functional Group (fig. 14) and Large-Bodied Species Abundances (fig. 15) having Procrustes scores similar to that of the highly distorted Callide Mine Large-Bodied Species Abundance ordination (fig. 7). It is worth noting that Large-Bodied Species actually has a slightly lower Procrustes score at German Creek than at Callide Mine, however. Species Presence/Absence (fig. 16) has a substantially lower Procrustes score than these three, but even this is higher than the Procrustes scores of three of the four analyses at Callide Mine.

Procrustes NMDS Comparisons	**Genus Abundance**	**Functional Group Abundances**	**Large-Bodied Species Abundance**	**Species Presence/ Absence**
Callide	18.95	20.9	18.42	13.92

Table 14: *Procrustes analysis results for comparison of simplified analyses to NMDS ordination of ant species abundance at German Creek Mine. Lower Procrustes scores = greater similarity of ordination based on that approach to that produced by species abundance approach.*

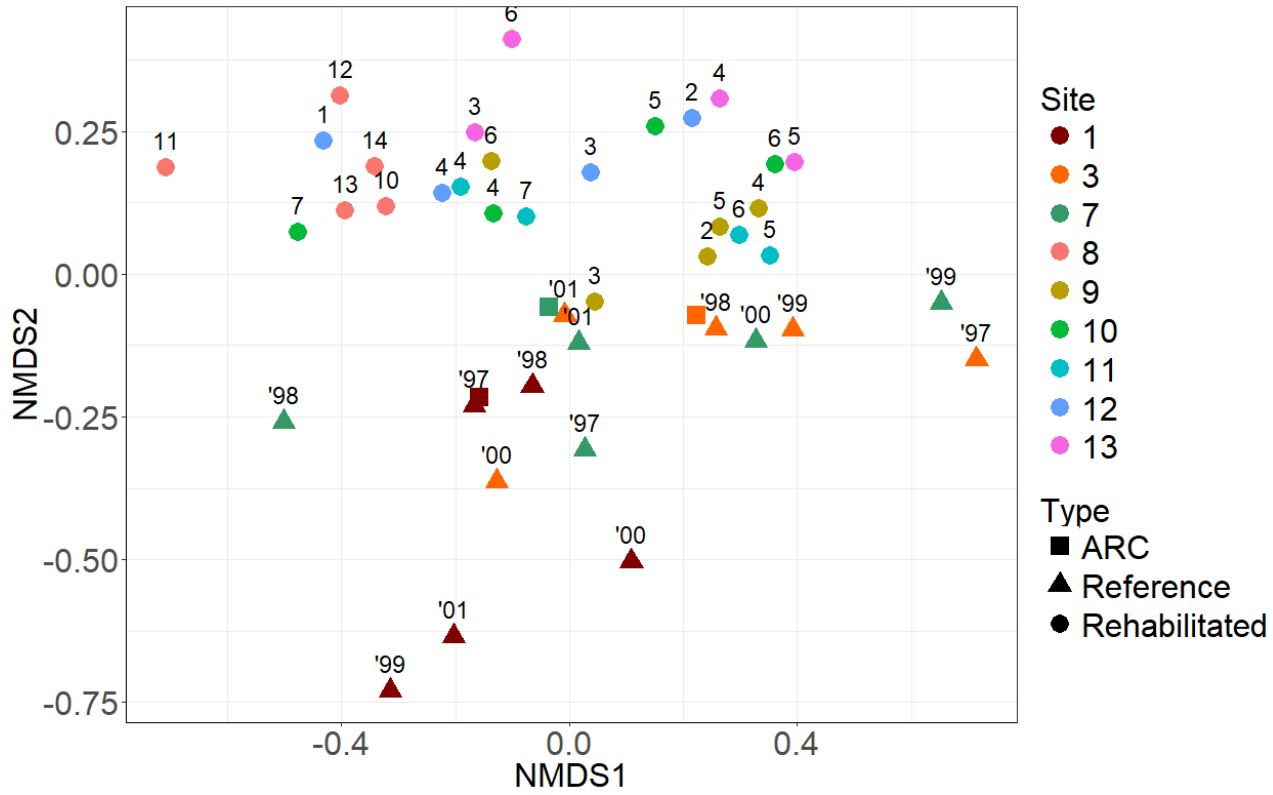


Figure 13: *Non-metric multidimensional scaling ordination of German Creek Mine Sites based on ant genus abundances, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (1997-2001) for reference site samples. Two-dimensional stress = 0.1718107*

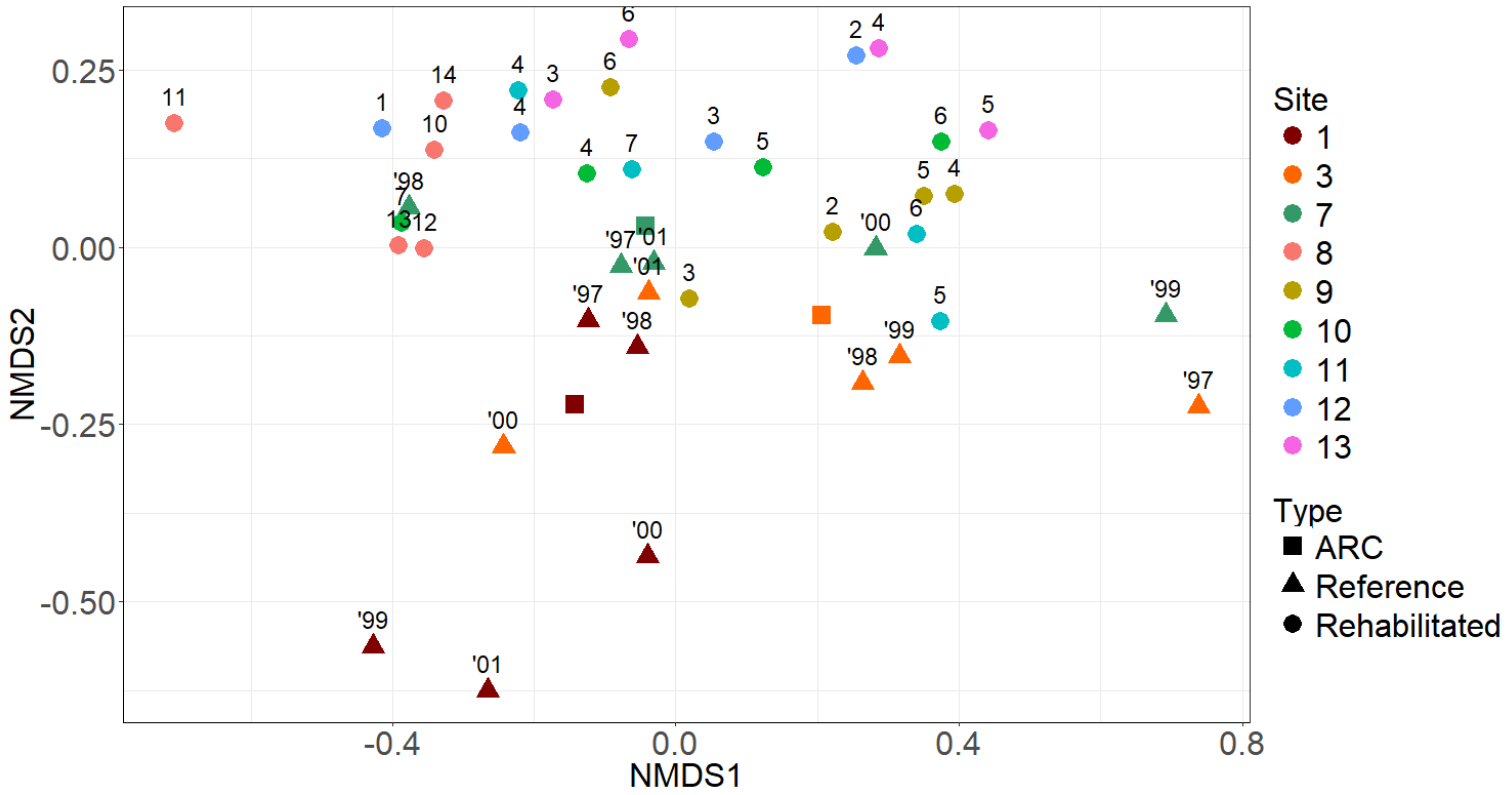


Figure 14: *Non-metric multidimensional scaling ordination of German Creek Mine Sites based on ant functional group abundances, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (1997-2001) for reference site samples. Two-dimensional stress = 0.1476129*

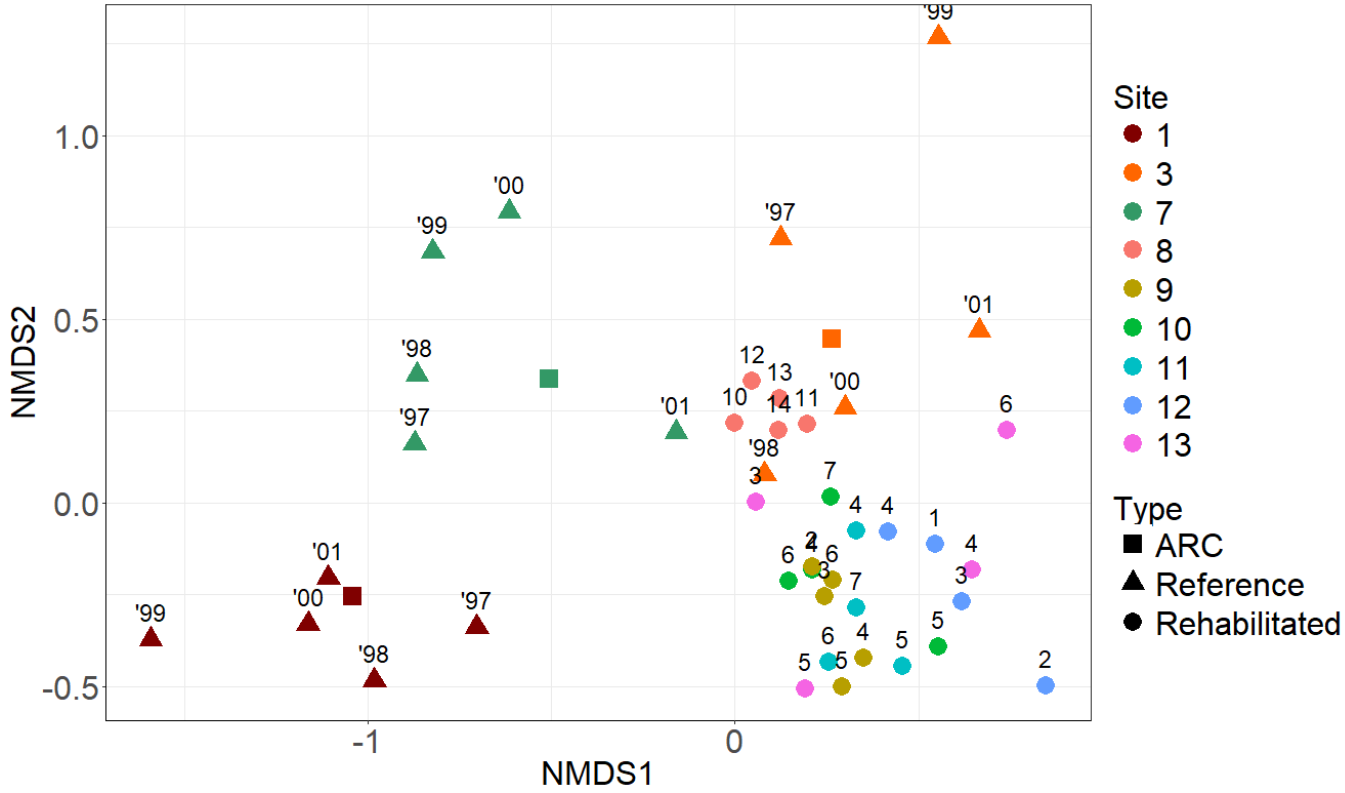


Figure 15: *Non-metric multidimensional scaling ordination of German Creek Mine Sites based on large bodied ant species abundances, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (1997-2001) for reference site samples. Two-dimensional stress = 0.174613*

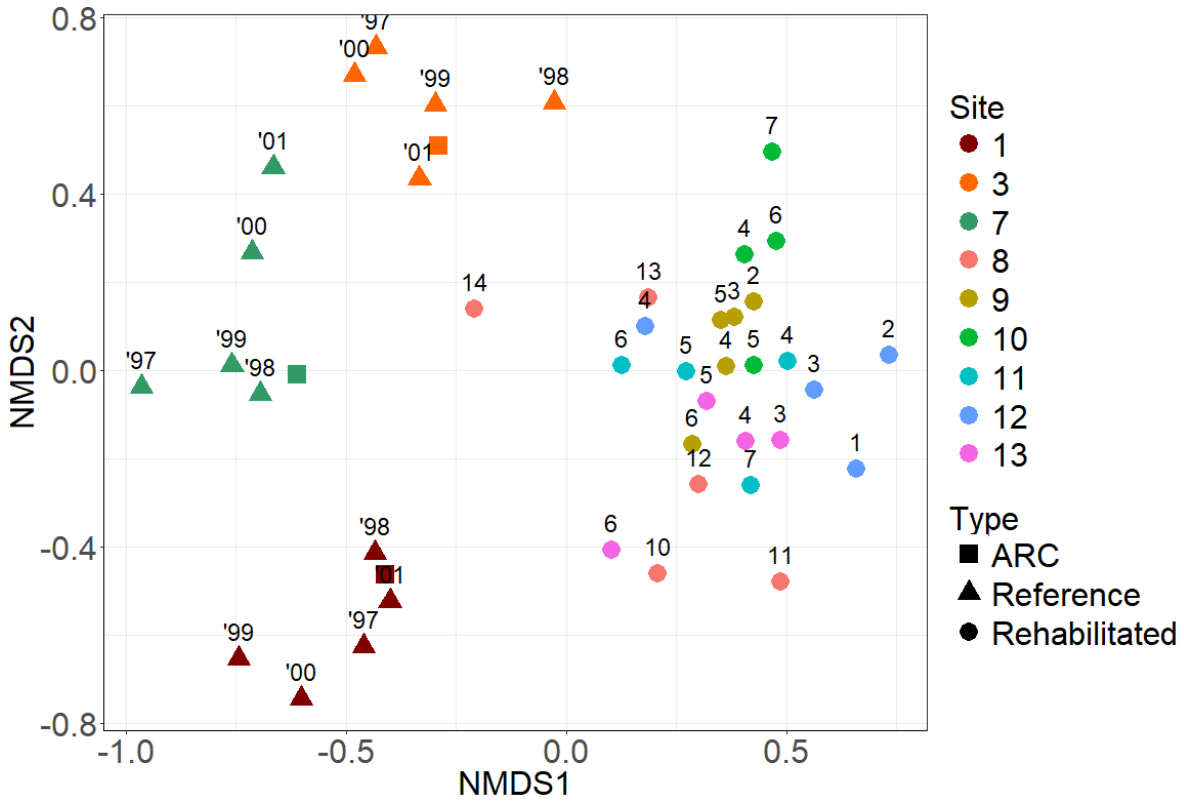


Figure 16: *Non-metric multidimensional scaling ordination of German Creek Mine Sites based on ant species presence and absence, comparing reference (triangle) and rehabilitated (circle) site types. ARCs (square) are also included to indicate their positioning within reference clusters. Point labels indicate age since rehabilitation (years) of rehabilitated site samples, or year sample was taken (1997-2001) for reference site samples. Two-dimensional stress = 0.2101377*

These results are reflected by the results of Kappa analysis of convergence/divergence patterns in ARC-Dissimilarity results (table 15), with only Species Presence/Absence obtaining an agreement with Species Abundance score which was significantly different from random classification of sites, or moderately reliable. This in turn corresponds to a high percentage agreement of 77.78%

Convergence/ Divergence	**Genus Abundance**	**Functional Group Abundance**	**Large-Bodied Species Abundance**	**Species Presence/ Absence**
Kappa	0.333	0.365	-0.169	0.532*
% Matching	66.67	66.67	44.44	77.78

Table 15: *Cohen's Kappa values of the four simplified analyses for matching site classification as converging or diverging from the ARC with results of Species Abundance analysis at German Creek Mine. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001)*

None of the four simplified analyses obtained significant intraclass correlation coefficients of agreement with Species Abundance analysis for estimated times to convergence (table 16), and ICCs were consistently low across the board. ICCS for the four simplified analyses were very similar.

EYCA	Genus Abundance		Functional Group Abundances		Large-Bodied Species Abundance		Species Presence/ Absence	
	ICC	Confidence Interval	ICC	Confidence Interval	ICC	Confidence Interval	ICC	Confidence Interval
Scores	0.038	-0.449-0.493	0.138	-0.303-0.547	0.103	-0.375-0.536	0.073	-0.371-0.503

Table 16: *Intraclass Correlation Coefficients (ICC) of the four simplified analyses at Callide Mine for matching the estimated years to convergence with ARCs (EYCA) of each reference site with results of Species Abundance analysis at German Creek Mine. Colours indicate reliability (Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = ≤0.01, *** = ≤0.001).*

The three analyses, Procrustes, Convergence/Divergence Kappa and EYCA ICC, present a more consistent picture of which of the simplified analyses would be appropriate surrogates than at Callide Mine, although this is largely due to the far higher number of non-significant results – Species Presence/Absence is portrayed by both Procrustes and Kappa analyses as having substantially greater fidelity to Species Abundance analysis than the other three simplified analyses. When the rankings and scores of the four analyses are considered in detail, there is little agreement between the three analyses. Procrustes analysis scores the Large-Bodied Species, Genus and Functional Group Abundance analyses as similarly distant from the Species Abundance analysis NMDS, while Kappa analysis ranks Large-Bodied Species Abundance as the weakest analysis by a substantial margin. ICC scores are more consistent with Procrustes, but score all four simplified analyses similarly, rather than rating Species Presence/Absence as having higher agreement. Indeed, it is not even first-ranked. What is consistent across the three analyses is that all four simplified analyses have lower agreement with Species Abundance analysis at German Creek than at Callide Mine.

Section 2.2 – Genus Abundance

Genus Abundance analysis performed poorly (table 17), with a high Procrustes score and non-significant agreement with Species Abundance ARC-Dissimilarity results at both levels, convergence/divergence and EYCA.

	Procrustes	**Convergence/ Divergence**		**EYCA**	
Scores	18.65	0.333	66.67%	0.038	-0.449-0.493
Rank	3	3		4	

Table 17: *Summary of scores and ranks of Genus Abundance analysis at German Creek Mine in each of the three agreement analyses. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001).*

As with Callide Mine, Genus Abundance ARC-Dissimilarity patterns for German Creek (fig. 17) show much lower dissimilarities between rehabilitated and reference sites. Subsequently there is much greater overlap between the two, with the majority of rehabilitated site samples actually overlapping with reference site samples on the y-axis (dissimilarity from the ARC), even as some are clearly diverging from the reference sites. As with Callide Mine, there is a degree of homogenisation between the convergence patterns respective to the three reference sites, although unlike at Callide Mine they do not closely resemble the convergence patterns of any of the three reference sites under Species Abundance analysis. Furthermore, while the overall directions of trends are similar, the underlying positions of individual site samples appear much more varied between the three sets of convergence patterns than at Callide Mine.

Figure 17 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 1 (17a), 3 (17b) and 7 (17c) with site age since rehabilitation (years) at German Creek Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 17a: *Convergence with Site 1, German Creek Mine*

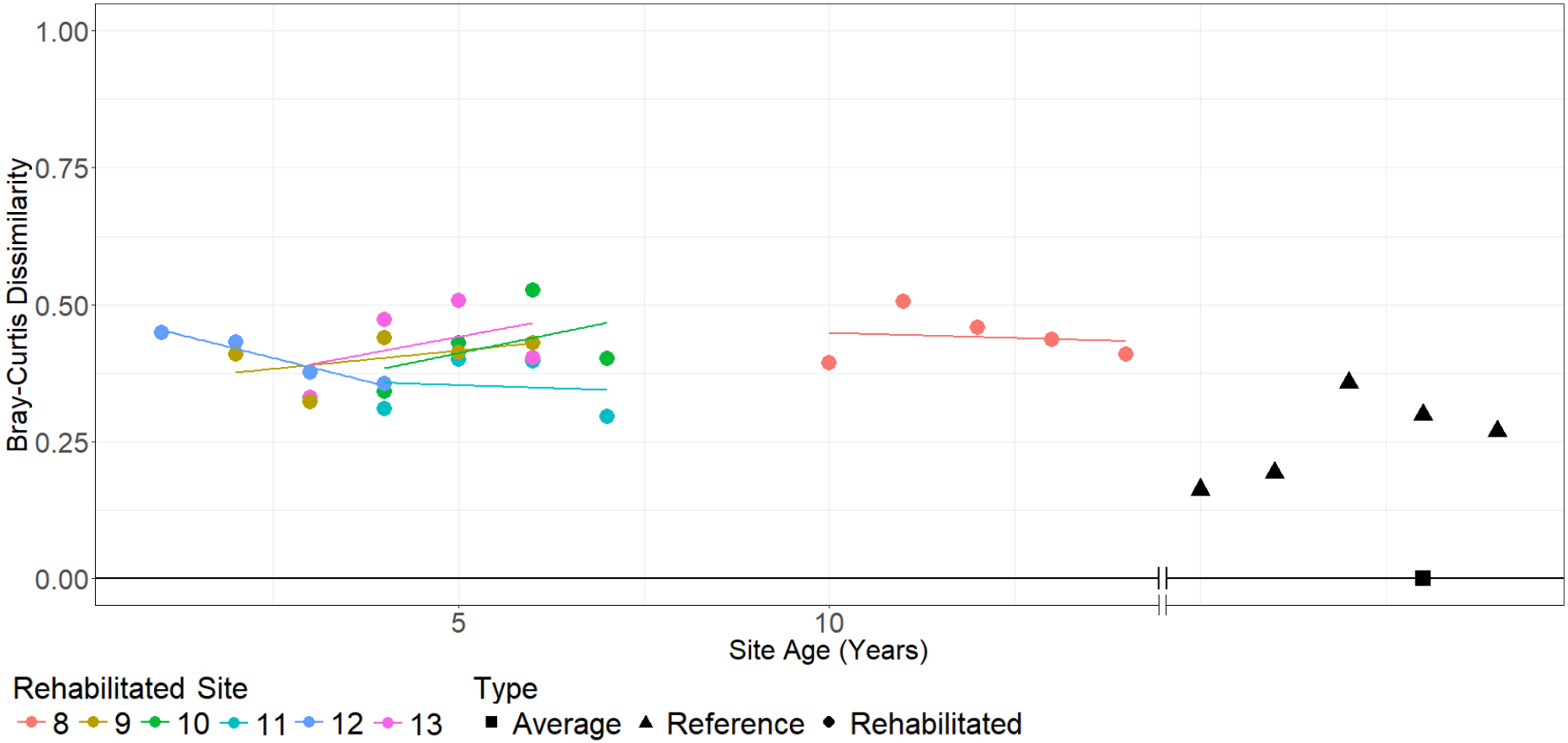


Fig 17b: *Convergence with Site 3, German Creek Mine*

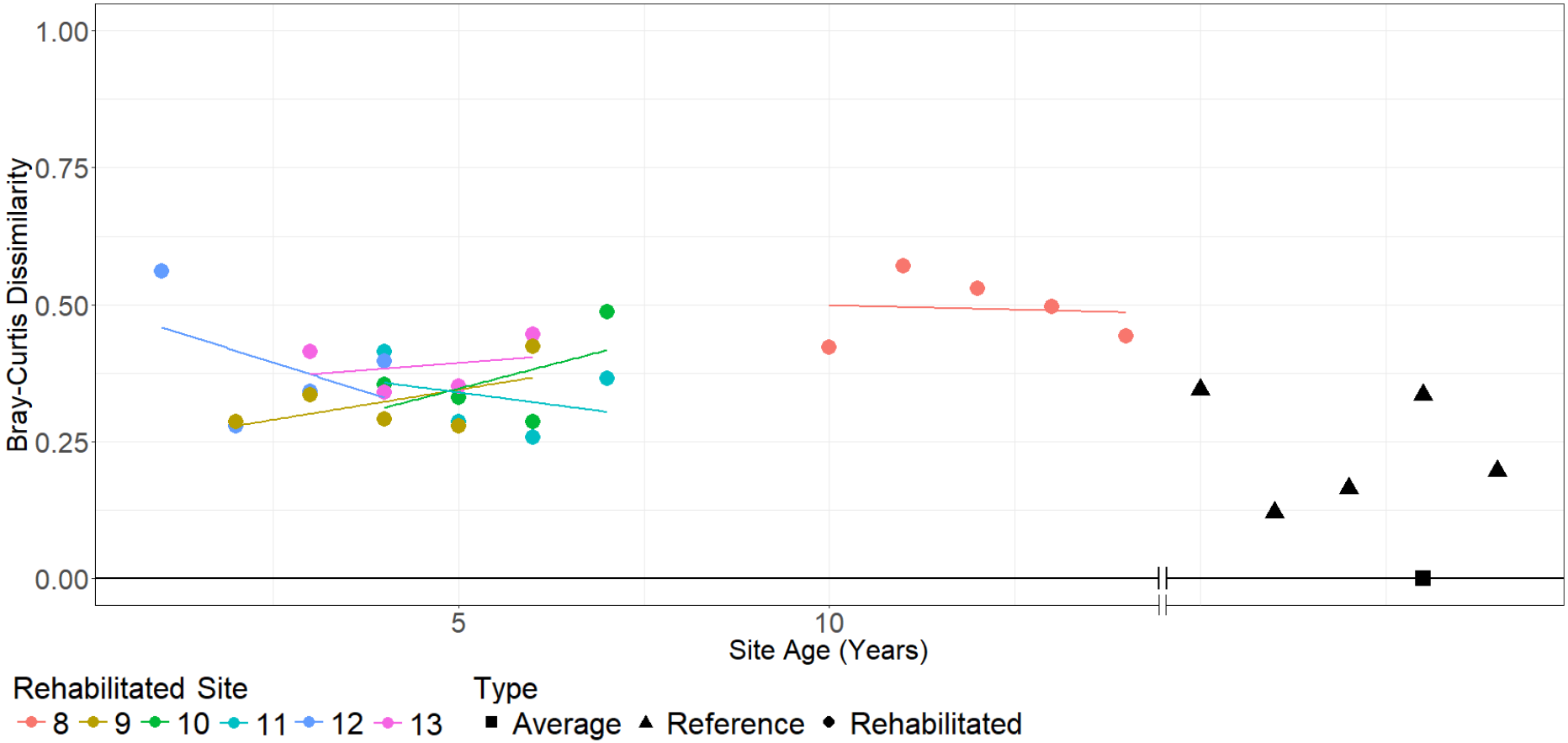
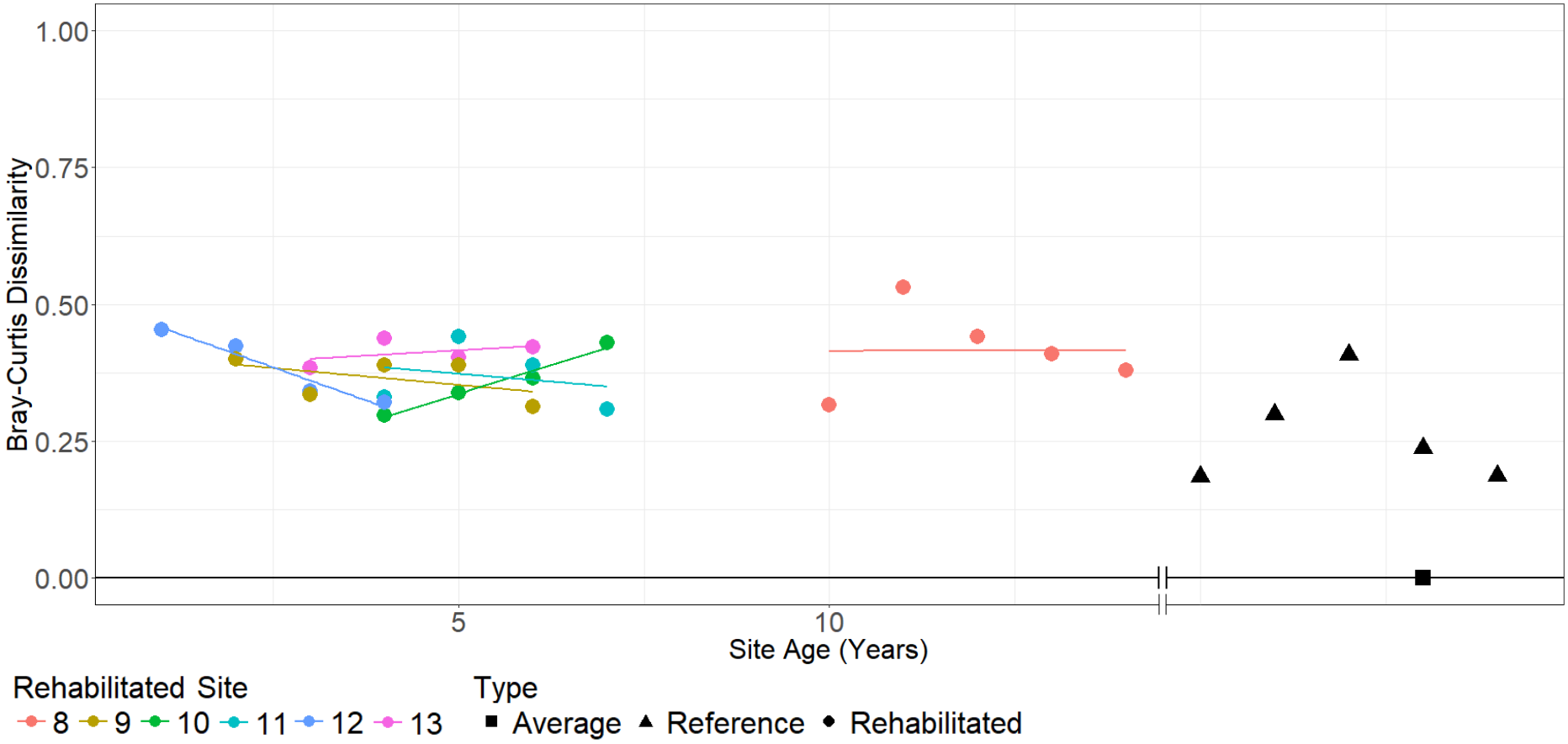


Fig 17c: *Convergence with Site 7, German Creek Mine*



In the NMDS ordination (fig. 13) for Genus Abundance, reference and rehabilitated sites remain separated along the primary axis of convergence (NMDS axis 2) with the exception of rehabilitated site 9 year 3. However, the gap between reference and rehabilitated sites is very small, and clearly smaller than the distance between the reference site ARCs and some samples from those reference sites, in congruence with the ARC-Dissimilarity results (fig. 17). The reference sites are also quite closely clustered together, as in the Genus Abundance analysis of Callide Mine data (fig. 5), explaining the increased homogenisation of results in the ARC-Dissimilarity results (fig. 17). Two additional notable differences between the ordination for Genus Abundance analysis and that of Species Abundance analysis are that references 3 & 7 in particular now have much more overlap, being completely separated in the Species Abundance analysis, and that the clustering of rehabilitated sites 9,10,11&12 is no longer apparent. There is also much more movement up and down the secondary axis (NMDS1) rather than the primary axis of convergence, with relatively little directly convergent movement towards the reference sites.

Section 2.3 – Functional Group Abundance

Functional Group Abundance analysis (table 13) performed as poorly as Genus Abundance analysis, with a high Procrustes score and non-significant agreement with Species Abundance ARC-Dissimilarity results at the levels of both convergence/divergence and EYCA.

	Procrustes	**Convergence/ Divergence**		**EYCA**	
Scores	20.9	0.365	66.67%	0.138	-0.303-0.547
Rank	4	2		1	

Table 18: *Summary of scores and ranks of Functional Group Abundance analysis at German Creek Mine in each of the three agreement analyses. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001).*

The ARC-Dissimilarity convergence patterns (fig. 18) reveal that, as at Callide Mine, the Bray-Curtis dissimilarity between rehabilitated and reference sites is even lower than under Genus Abundance analysis, with many rehabilitated sites occurring in the same range of dissimilarities as the reference sites for the entirety of the sampling period, particularly reference site 7 (fig.18c). However, the convergence patterns relative to the three reference sites are more differentiated than for Genus Abundance analysis.

Figure 18 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 1 (18a), 3 (18b) and 7 (18c) with site age since rehabilitation (years) at German Creek Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 18a: *Convergence with Site 1, German Creek Mine*

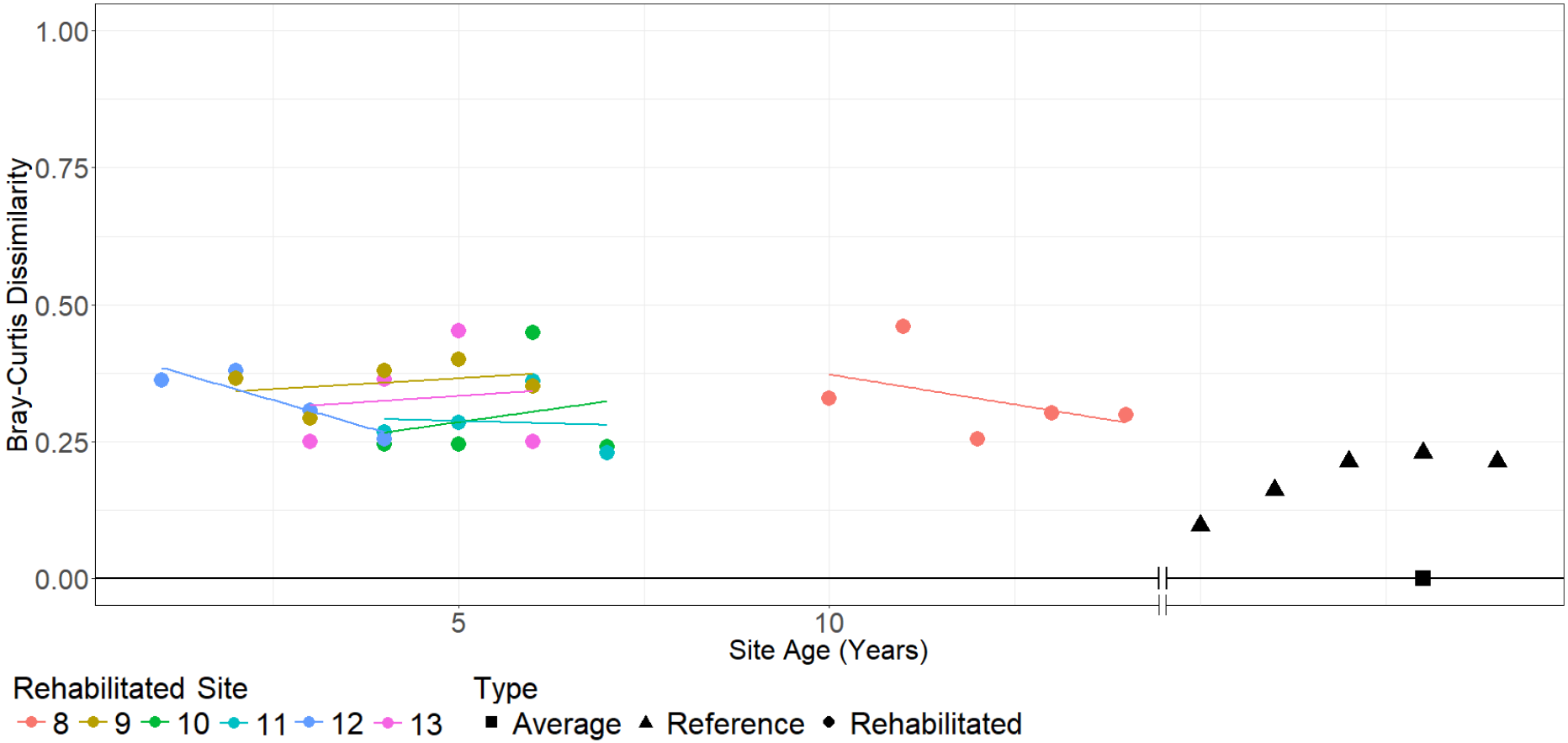


Fig 18b: *Convergence with Site 3, German Creek Mine*

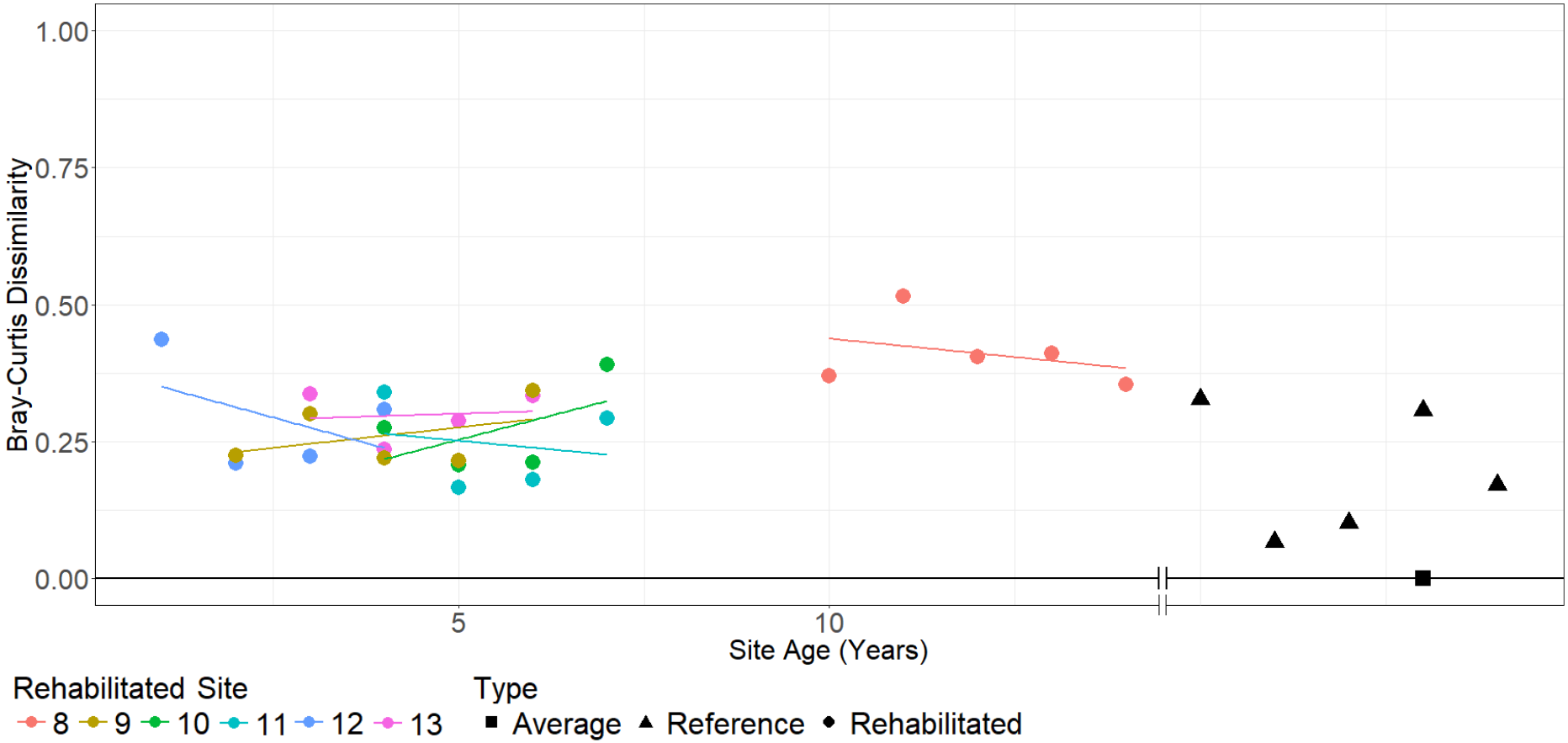
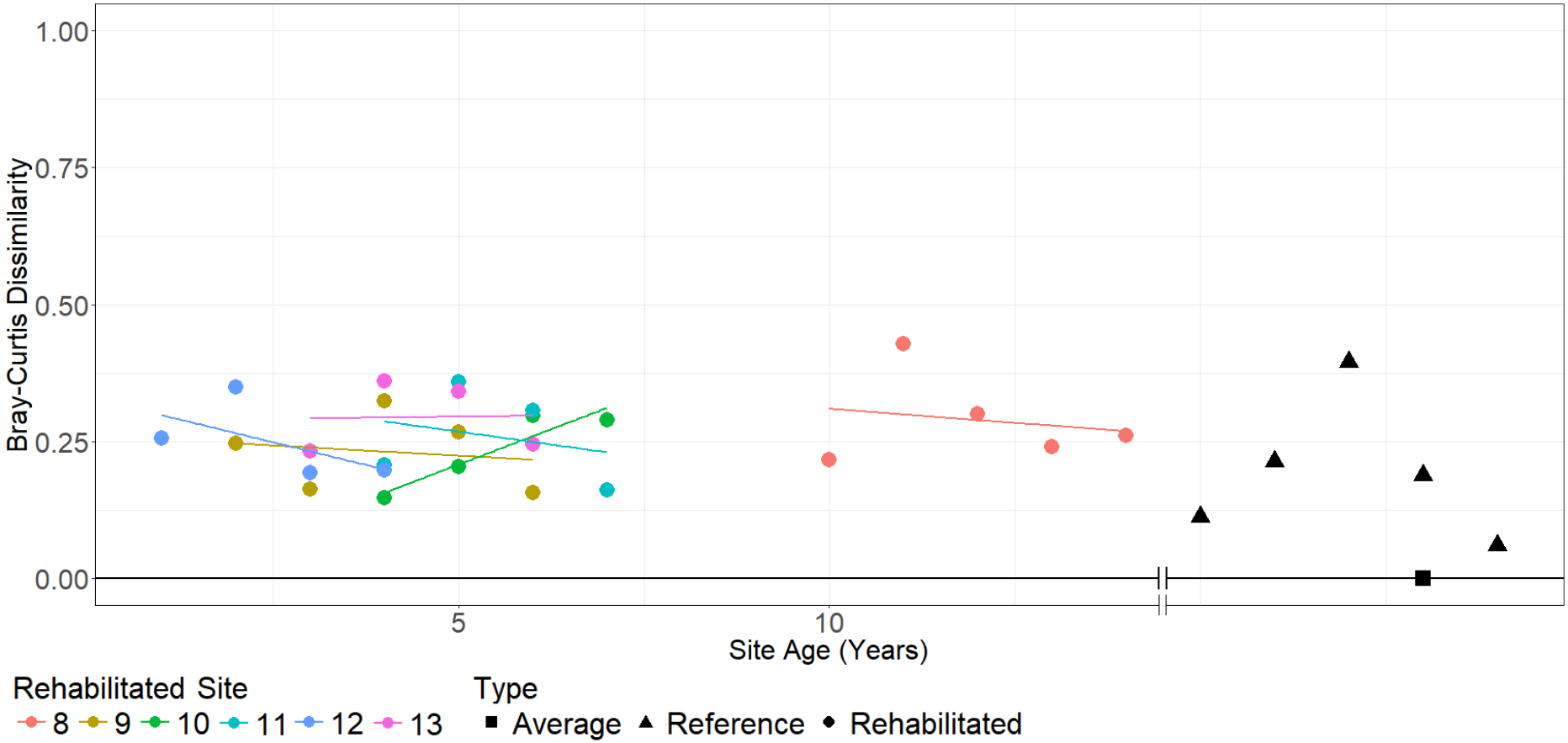


Fig 18c: *Convergence with Site 7, German Creek Mine*



The NMDS ordination for Functional Group Abundance analysis (fig. 14) is very similar to that of Genus Abundance analysis, with very similar consequences. The primary difference is that reference site 7 now overlaps with the rehabilitated sites on the primary axis of convergence (NMDS axis 2), corresponding with the ARC-Dissimilarity model for reference site 7 (fig. 18c). The whole ordination is also more condensed, with lower dissimilarities between sites, as in the ARC-Dissimilarity results. However, the reference site ARCs are also slightly further apart than under Genus Abundance analysis, reflecting the greater differentiation of ARC-Dissimilarity convergence patterns respective to each one.

Section 2.4 – Large-Bodied Species Abundance

Large-Bodied Species Abundance analysis (table 19) performed as poorly as Genus and Functional Group Abundance analyses, with a high Procrustes score and non-significant agreement with Species Abundance ARC-Dissimilarity results at the levels of both convergence/divergence and EYCA. It also experienced an even greater drop in raw percentage agreement at the convergence/divergence level, to below 50%. Despite a substantial drop in performance at both levels of ARC-Dissimilarity analysis, the Procrustes score is actually slightly better than at Callide Mine (Callide Procrustes score = 19.2).

	Procrustes	**Convergence/ Divergence**		**EYCA**	
Scores	18.42	-0.169	44.44%	0.103	-0.375-0.536
Rank	2	4		4	

Table 19: *Summary of scores and ranks of Large-Bodied Species Abundance analysis at German Creek Mine in each of the three agreement analyses. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001).*

Uniquely among the simplified analyses, in the ARC-Dissimilarity results for Large-Bodied Species Abundance at German Creek (fig. 19) most of the discrepancies between the Large-Bodied Species Abundance and Species Abundance analyses are changes in convergent/divergent classification of rehabilitated sites relative to each reference site, rather than more precise differences in estimated convergence time. Like at Callide Mine, these ARC-Dissimilarity results are not homogenous and also do not show a consistent degree of dissimilarity reduction across the convergence patterns, with the extent of reduction in initial Bray-Curtis dissimilarity of the rehabilitated sites relative to each reference site varying considerably.

Looking at the NMDS ordination for Large-Bodied Species Abundance analysis (fig. 15), we see a few major changes. The most important is the change in position of reference site 7, which has moved from the bottom of the ordination to the top. Another notable change is that reference site 3 now overlaps with all rehabilitated sites on axis 1, although not on axis 2. However, rehabilitated site 8 now overlaps with reference site 3, on both axes, rather than with the other rehabilitated sites, and appears to be relatively static, like TH91 at Callide Mine in the ordinations for Species Abundance and Species Presence/Absence analyses. This substantially differentiated ordination plot correlates with the radically different set of ARC-Dissimilarity results.

Figure 19 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 1 (19a), 3 (19b) and 7 (19c) with site age since rehabilitation (years) at German Creek Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 19a: *Convergence with Site 1, German Creek Mine*

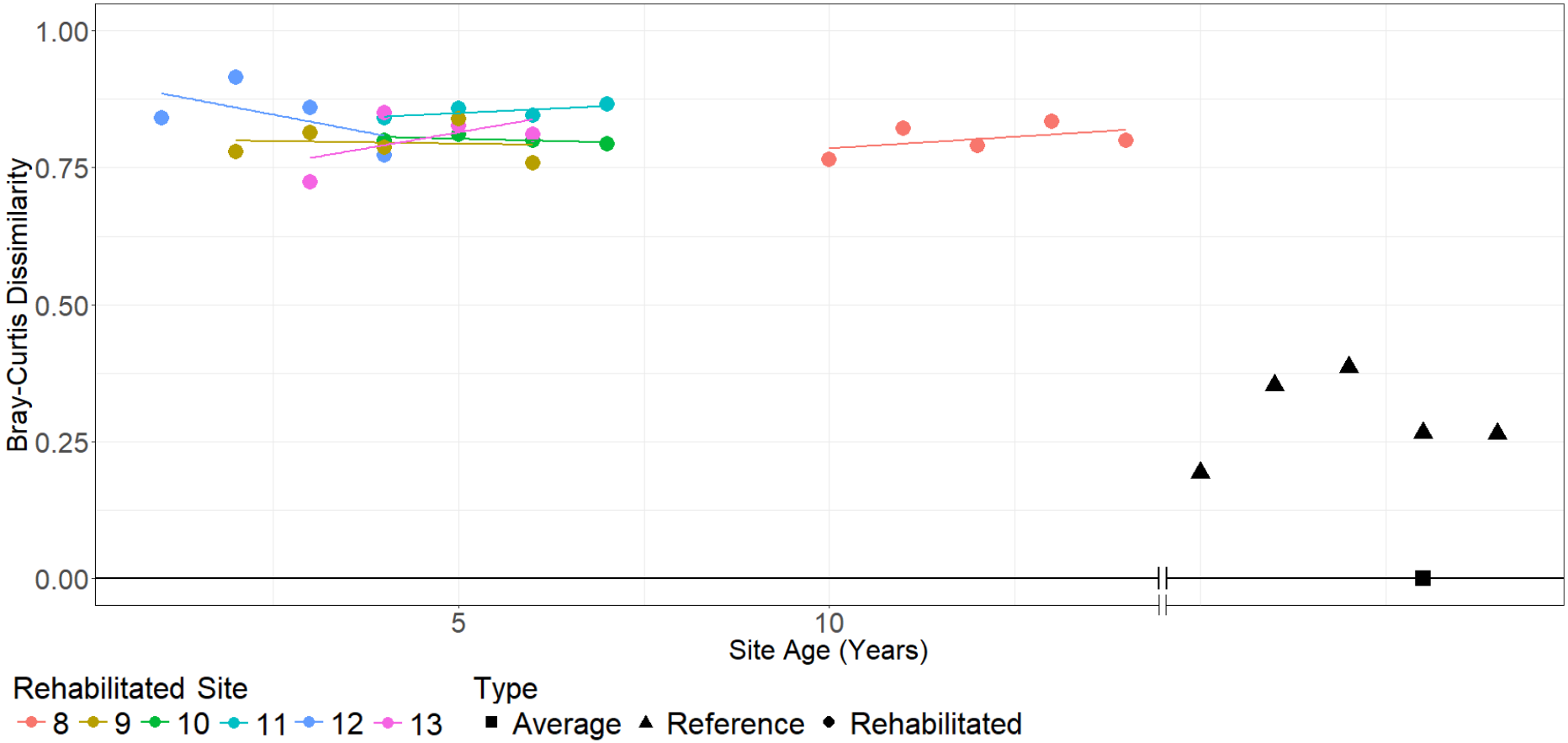


Fig 19b: *Convergence with Site 3, German Creek Mine*

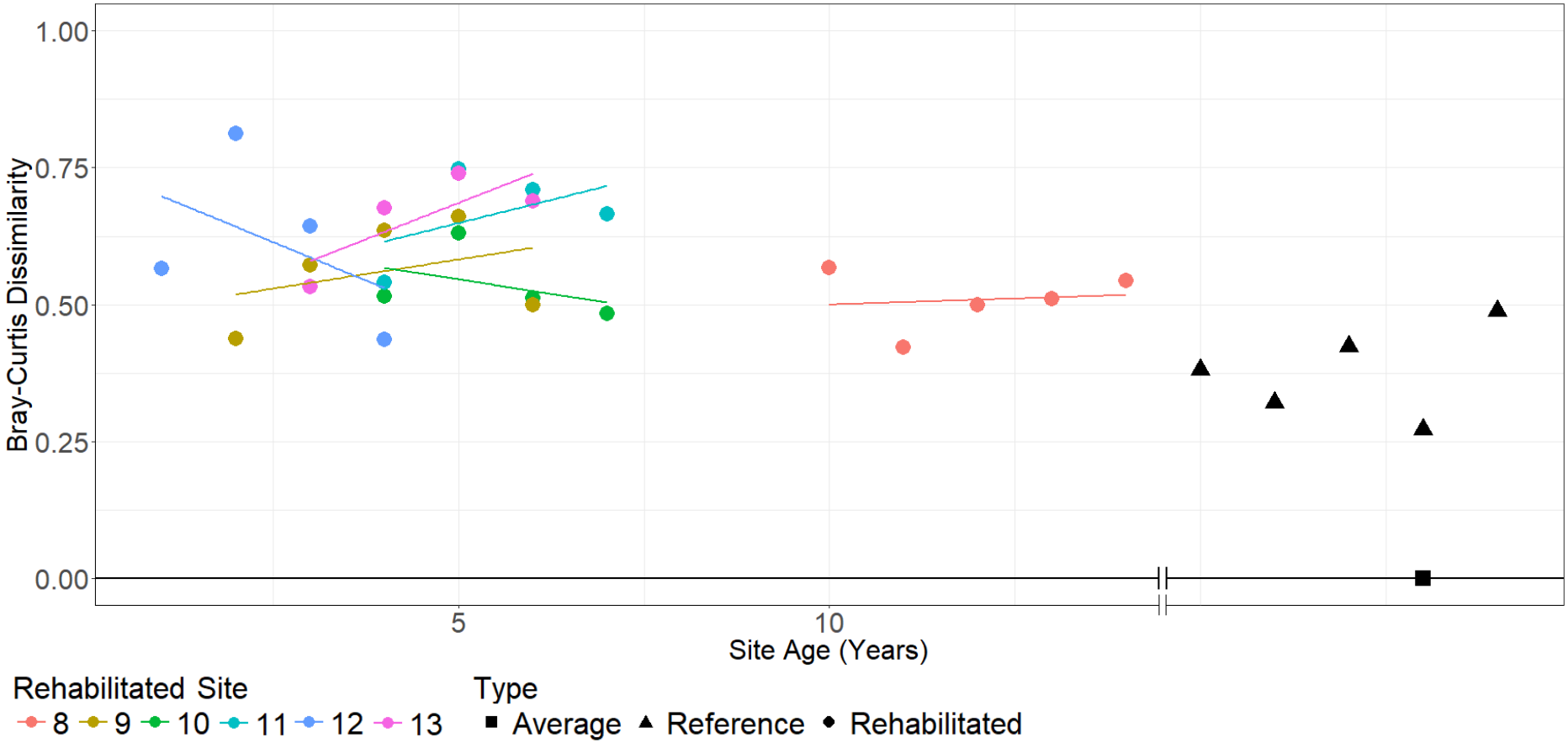
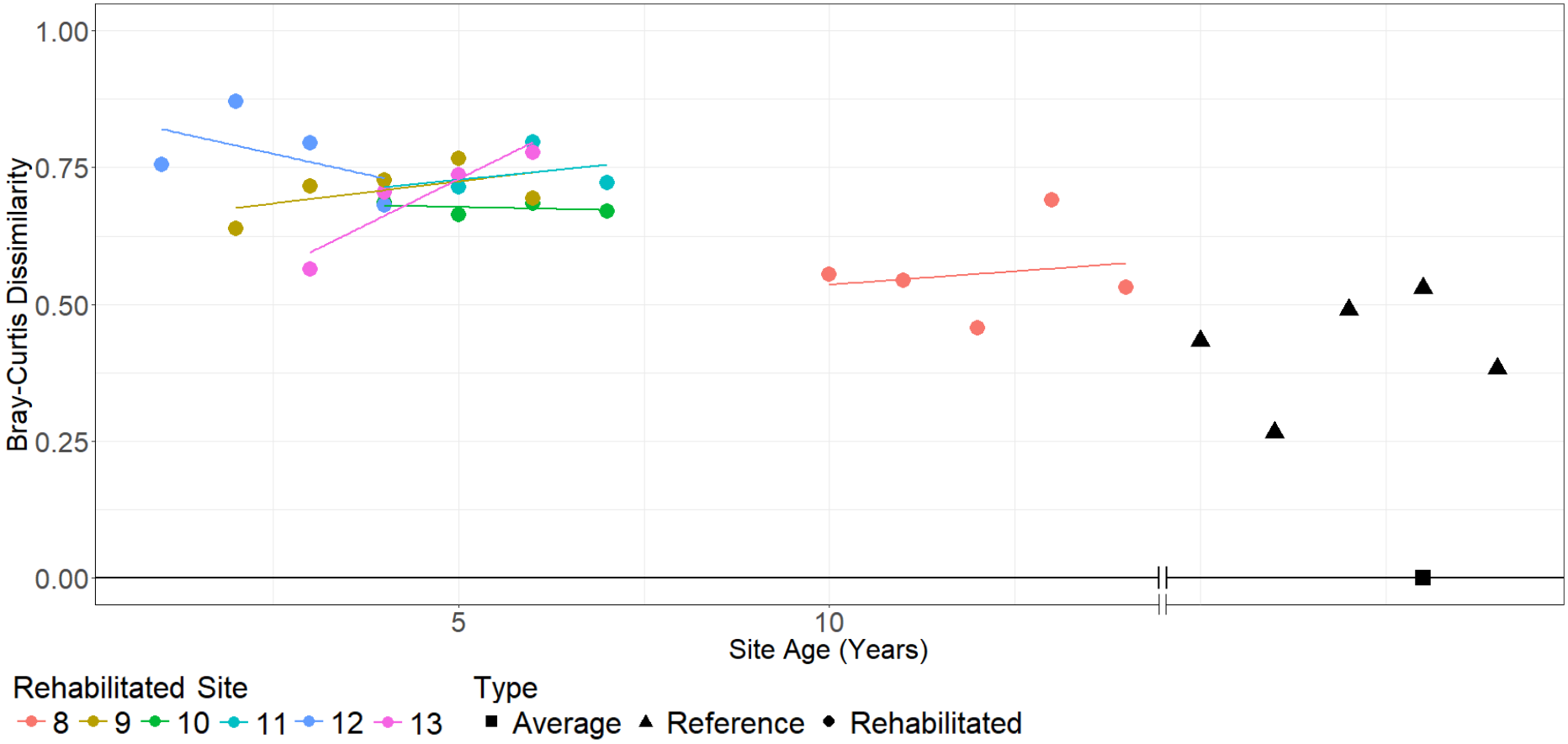


Fig 19c: *Convergence with Site 7, German Creek Mine*



2.5 – Species Presence/Absence

Uniquely among the four simplified analyses, Species Presence/Absence has a significant level of agreement at German Creek with Species Abundance analysis at the convergence/divergence level, achieving a similar level of agreement as at Callide Mine, although not quite as high (Callide: Kappa = 0.571). It also has the lowest Procrustes scores by a substantial margin. However, like the other simplified analyses it does not achieve significant agreement with Species Abundance analysis on estimated years to convergence.

	Procrustes	**Convergence/ Divergence**		**EYCA**	
Scores	13.92	0.532*	77.78%	0.073	0.371-0.503
Rank	1	1		1	

Table 20: *Summary of scores and ranks of Species Presence/Absence analysis at German Creek Mine in each of the three agreement analyses. Colours indicate reliability (Dark Grey = No Agreement, Light Grey = Extremely Poor Reliability, Red = Poor Reliability, Yellow = Moderate Reliability, Green = Good Reliability, Blue = Excellent Reliability) & asterisks indicate level of significance (= >0.05, . = 0.05, * = <0.05, ** = <0.01, *** = <0.001).*

The ARC-Dissimilarity results for Species Presence/Absence reveal a slight decrease in initial Bray-Curtis dissimilarity of rehabilitated sites from reference site ARCs (fig. 20) compared to Species Abundance, but rehabilitated sites do not overlap with reference sites under this analysis, despite there being an overlap with reference sites 3 & 7 in Species Abundance analysis (fig. 4b & 4c (Chapter 2 Results)), and changes in dissimilarity relative to reference site 3 are notably reduced (20b) from Species Abundance analysis, suggesting Species Presence/Absence analysis may actually be more conservative than Species Abundance. As at Callide Mine, convergence patterns relative to each reference site are not homogeneous and resemble their Species Abundance counterparts.

Figure 20 (next page): *Change in Bray-Curtis Dissimilarity of Rehabilitated Mine Sites from ARCs of Reference Sites 1 (20a), 3 (20b) and 7 (20c) with site age since rehabilitation (years) at German Creek Mine. These figures show the dissimilarity of the rehabilitated sites (coloured circles), relative to the ARC (black square) of the reference site, against site age since rehabilitation. The reference site's samples' (black triangles) dissimilarities from their ARC are included for comparison.*

Fig 20a: *Convergence with Site 1, German Creek Mine*

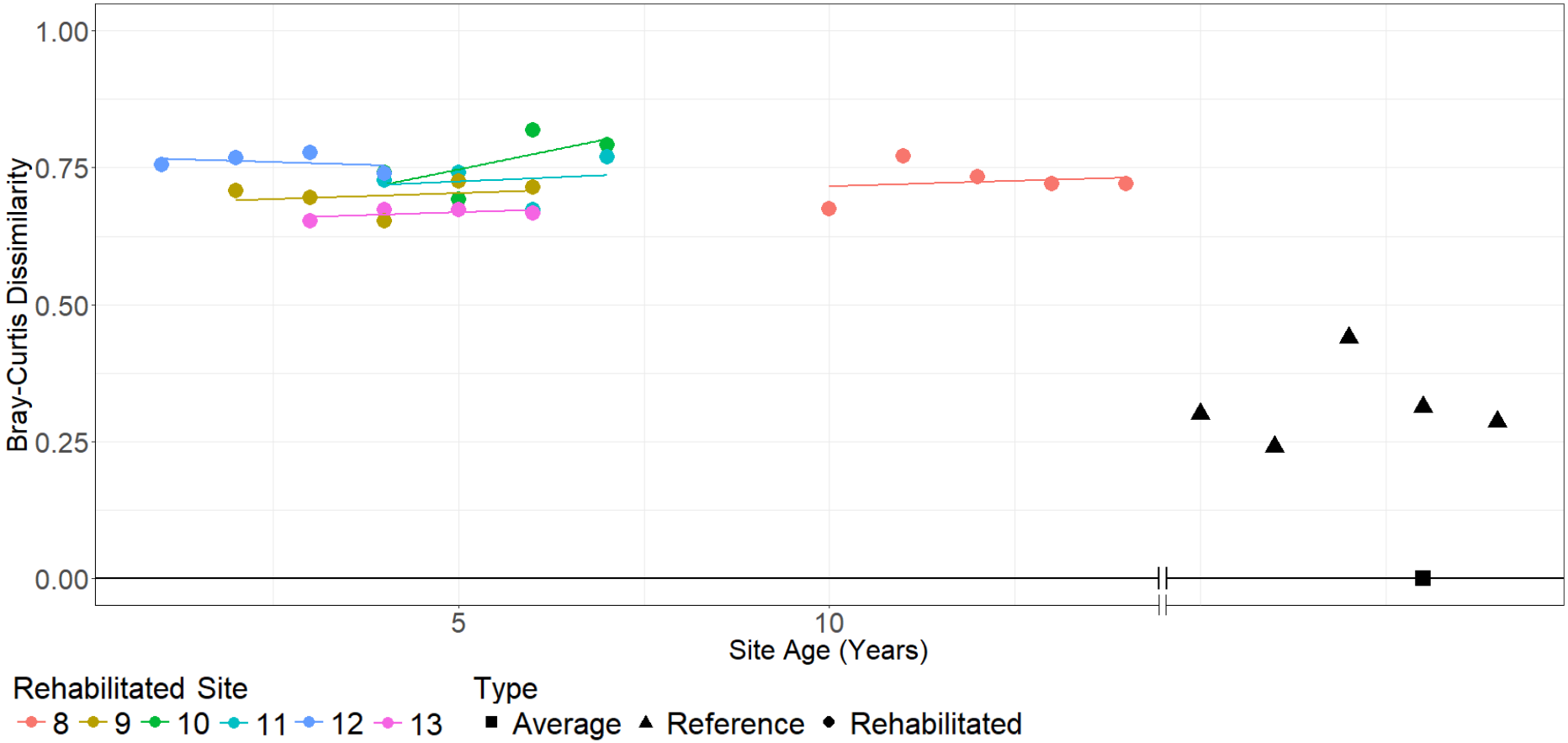


Fig 20b: *Convergence with Site 3, German Creek Mine*

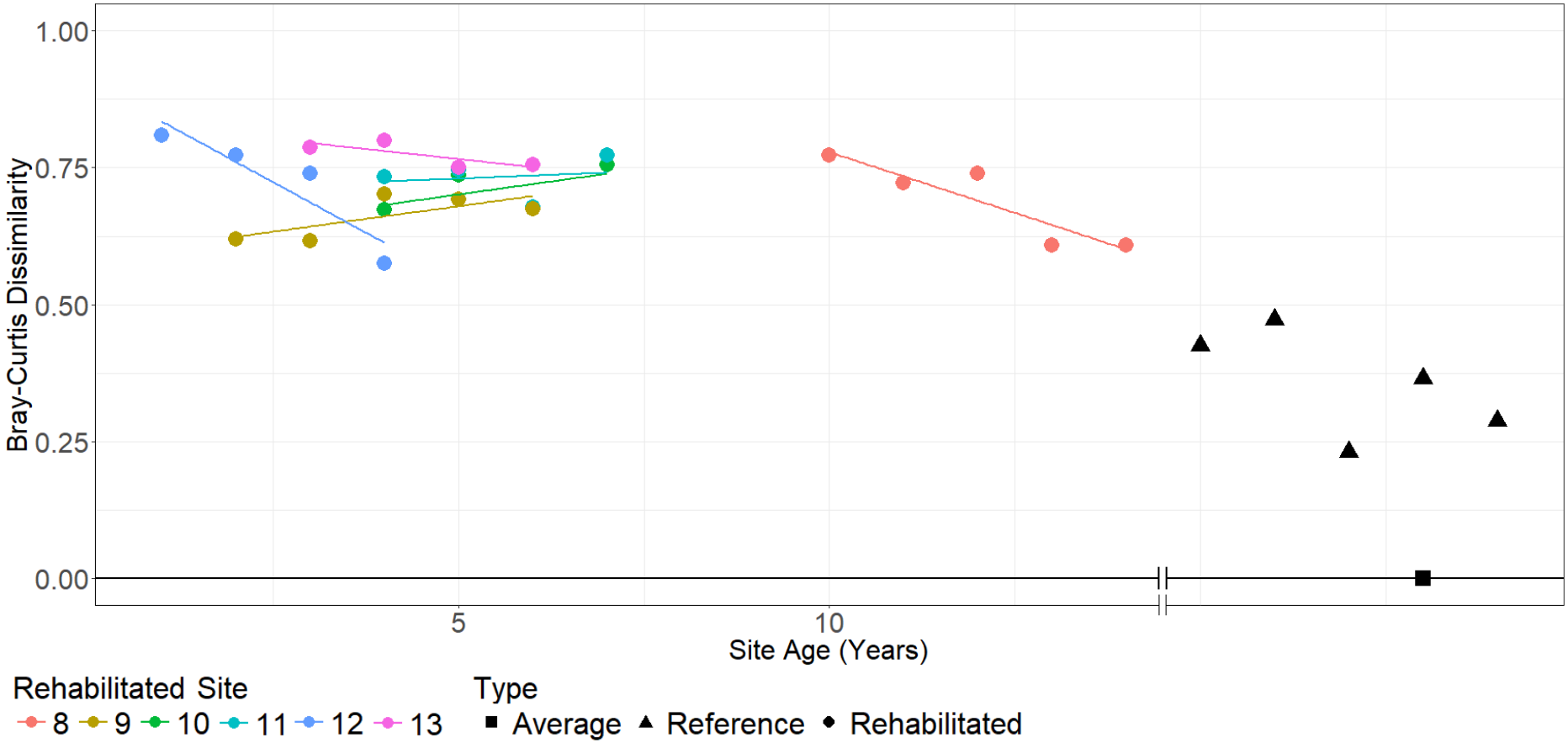
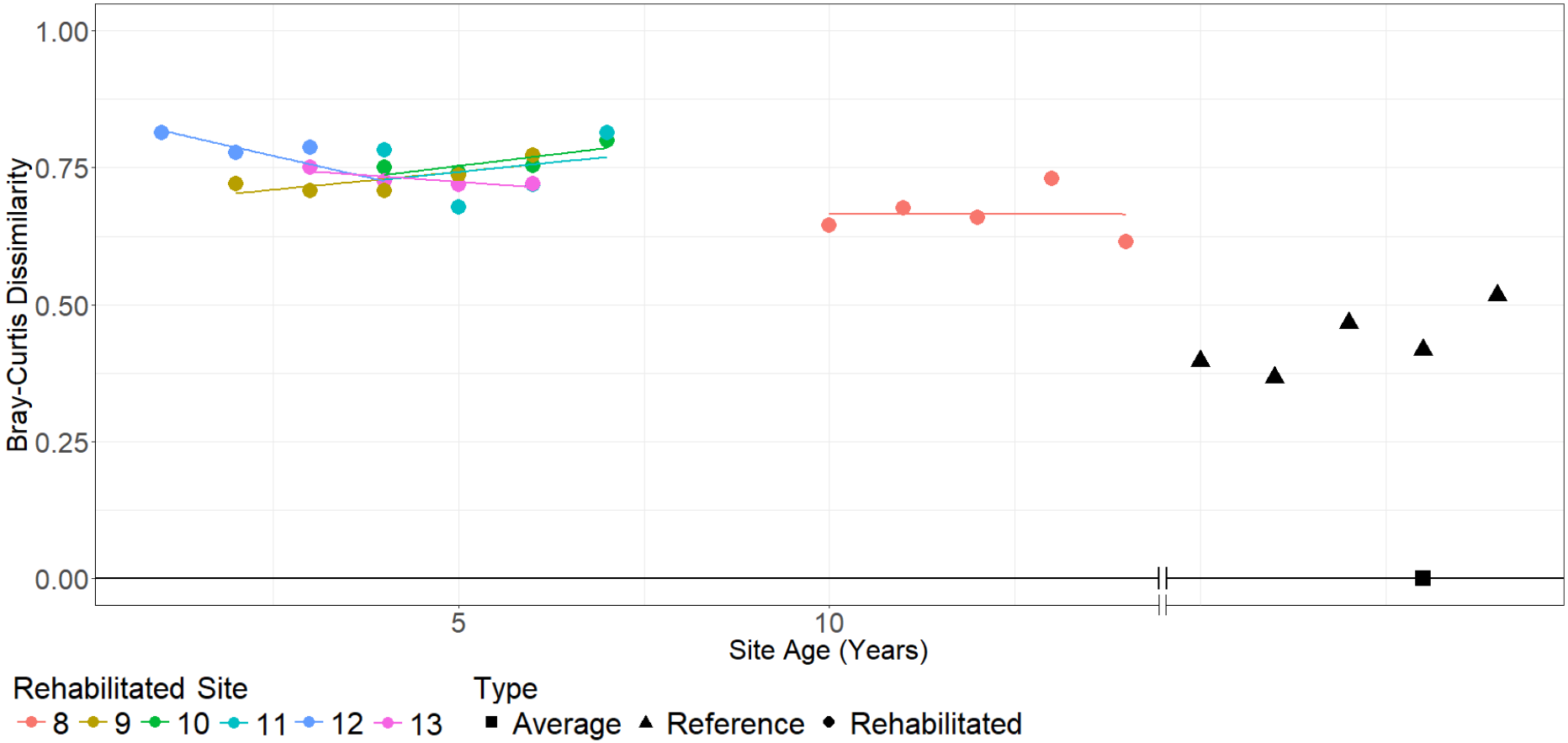


Fig 20c: *Convergence with Site 7, German Creek Mine*



This interpretation is supported by the NMDS ordination for Species Presence/Absence analysis, in which the reference and rehabilitated sites are more clearly separated with very little overlap. Like with Large-Bodied Species Abundance analysis, reference site 7 is shown to have moved from the bottom to the top of the ordination. The cluster of sites 9, 10, 11 & 12 is also absent from this ordination, and they all appear to be less divergent – indeed site 9 appears to be slightly convergent, and the convergent trends of sites 8, 12 & 13 are much less ambiguous, with corresponding increase in number of reference sites converged on for sites 12 & 13, although site 8 is now considered divergent from reference site 1. Despite having the most similar ordination in Procrustes analysis to Species Abundance (table 14) this ordination still represents a notable departure from that of Species Abundance, with much less ambiguous or erratic movement of sites towards or away from reference sites than in Species Abundance. Interestingly, unlike Large-Bodied Species Abundance analysis, these substantial deviations do not have correspondingly substantial deviations in the ARC-Dissimilarity results for Species Presence/Absence analysis, which are the most similar to those Species Abundance at the convergence/divergence level of the four simplified analyses at German Creek.

2.6 – Rehabilitated Site Trends (NMDS Ordination)

In qualitative interpretation of NMDS ordinations (table 21) of each of the simplified analyses, none of the four were able to replicate the trends of rehabilitated sites seen in the ordination for Species Abundance analysis, with the exception of for rehabilitated site 12. The major flaw in all four analyses was that the "clustering" of sites 9, 10, 11 & 12 in the original analysis is not maintained, with a loose association between these sites being present at best and substantially different patterns for each site apart from site 12, which maintained its overall convergent trajectory. Another major change in the ordinations of Large-Bodied Species Abundance and Species Presence/Absence analyses is the shifting position of reference site 7, from the bottom of the ordination to the top. As such, all four analyses undergo major shifts in the patterns of rehabilitated and reference sites, and do not maintain trends and patterns of rehabilitated site movement seen in the Species Abundance analysis.

Table 21 (next page): *Qualitative interpretation of rehabilitated site community trends in terms of convergence/divergence with reference sites at German Creek Mine, based on NMDS ordination for each analysis.*

	Original Analysis	*Simplified analyses*			
Rehabilitated Sites	***Species Abundance (fig. 3 (Chapter 2 Results)***	***Genus Abundance (fig. 13)***	***Functional Group Abundance (fig. 14)***	***Large-Bodied Species Abundance Species Abundance (fig. 15)***	***Species Presence/Absence (fig. 16)***
8	**Convergent on reference site 3, divergent from reference sites 1 & 7**	Erratic, little net movement and no convergence with reference sites	Erratic, little net movement and no convergence with reference sites	No net movement, stable position within space occupied by reference site 3	Convergent on reference sites 3 & 7, divergent from reference site 1
9	**Little movement until 2001, when rapidly diverged from reference sites**	Erratic, converges in 1997-98, but net divergence from reference sites	Erratic, converges in 1997-98, but net divergence from reference sites	Erratic, no net convergence	Erratic, slow convergence on reference sites – possibly divergence from reference site 3

10	**Little movement until 2001, when rapidly diverged from reference sites**	Erratic, diverging from reference sites but not on primary convergence axis	Erratic, diverging from reference sites but not on primary convergence axis	Net convergence with reference site 3 on secondary axis, no net convergence with sites 1 or 7.	No convergence on reference sites
11	**Little movement until 2001, when rapidly diverged from reference sites**	Erratic, little net movement.	Erratic, converges to point of overlap with reference sites in 1999 before diverging again, net movement is convergent with reference sites.	Net divergence with reference site 3 on secondary axis, no net convergence with sites 1 or 7.	Convergent in 1998-2000 but then massive divergence in 2001, so little net movement

12	**Erratic, net convergence with reference sites**	Erratic, net gradual convergence with reference sites	Erratic, little net movement	Little net movement	Relatively consistent convergence on reference sites
13	**Little movement, net convergence with reference sites**	Erratic, converges in 1999-2000 but net divergence from reference sites	Erratic, net movement is slight divergence from reference sites.	Erratic, net divergence from reference sites – no change relative to reference site 3.	Consistent convergence with reference sites, particularly reference site 1 – possibly divergent from reference site 3

2.7 – Summary (German Creek Mine)

None of the four simplified analyses reliably agreed with the results of the Species Abundance analysis of trends at German Creek Mine. Genus, Functional Group and Large-Bodied Species Abundance analyses did not achieve significant ARC-Dissimilarity agreement scores at either convergence/divergence or EYCA levels, and all three scored poorly in Procrustes analysis. Species Presence/Absence was more reliable, achieving moderate agreement reliability at the convergence/divergence level and a decent Procrustes score, although it did not achieve any significant agreement at the EYCA level either. All four simplified analyses likewise performed poorly in the qualitative assessment of NMDS ordinations, with key features of the Species Abundance ordination not retained in any of the four analyses, reflecting relatively high Procrustes analysis sums-of-squares scores in comparison to Callide Mine. Overall the performance of the four simplified analyses was substantially worse than at Callide Mine, with all but Species Presence/Absence ineffective at all levels of analysis.

3 – Conclusion

The key finding from this result is that the use of simplified analyses as a surrogate for Species Abundance is limited and highly contextual at best. First, the contextual limitation: there is the wide discrepancy in results between Callide Mine and German Creek, with two of the four simplified analyses (Genus and Functional Group Abundances) performing substantially worse in both NMDS and ARC-Dissimilarity, and Large-Bodied Abundance performing substantially worse in ARC-Dissimilarity. These results suggesting certain ant communities are more amenable to the use of surrogate taxonomic measures than others, and, when the relative consistency of Species Presence/Absence is considered, that the contextual requirements of each of the four analyses also differ.

Second, even the relative success of the four simplified analyses at Callide Mine is limited. While qualitative examination of NMDS ordinations and ARC-

Dissimilarity graphs at Callide Mine suggest broad patterns in community change can be conserved, statistical analysis of agreement between these simplified analyses and Species Abundance analysis reveals only moderate agreement reliability of ARC-Dissimilarity results even at the level of which direction communities are moving, converging or diverging, and non-significant agreement reliability for estimated convergence times, making them unreliable for use by practitioners. That said, moderate agreement reliability in Kappa analysis corresponds to over 80% agreement in practice for all four analyses, so they may in fact be more reliable at the broad scale of testing for convergence or divergence than these results suggest.

These results therefore provide scope for improvement of the use of simplified analyses by identifying potential systematic issues underpinning mismatches between these analyses and a full Species Abundance analysis.

Another important take-away is that the best simplified analysis to employ depends on the assessment method used. There is no conclusive agreement on the "best analysis" between Procrustes analysis of NMDS ordinations, ARC-Dissimilarity and qualitative interpretation of NMDS patterns, although Species Presence/Absence analysis does emerge as a potential "all-rounder" which, while not necessarily the best analysis under all of the three approaches, is consistent in its reliability across assessment methods and, to a lesser extent, across sites, as the only simplified analysis which is usable at any level of assessment across both sites.

Finally, these comparisons lend weight to the ARC-Dissimilarity approach as a supplementary means of assessing rehabilitated mine-site recovery, as the approach shows good correspondence to NMDS ordination patterns from the same mine under the same simplified analysis and is robust to datasets where the primary axis of convergence is not the axis of greatest variation, as discussed in chapter 2 (Chapter 2 Discussion, section 2).

Discussion

1 – Critical Discussion of Comparative Assessment Methods

The various analyses used here to assess the simplified analyses are effective means of drawing out patterns of convergence and recovery of rehabilitated sites, but also have a number of limitations it is important to acknowledge and understand. In this section I discuss these limitations and the contributions of these methods to understanding the patterns of convergence and recovery.

1.1 – ARC-Dissimilarity Comparison

We made use of a novel approach, ARC-Dissimilarity assessment, to assess whether sites were converging or diverging from reference sites by direct comparison of changes in dissimilarity between rehabilitated sites and the reference sites (see Chapter 2 Methods, section 2). Evaluating the fidelity of ARC-Dissimilarity results of the simplified analyses to those of Species Abundance analysis by comparing estimates of years to convergence has some limitations and drawbacks.

The ARC-Dissimilarity method uses linear models to predict the number of years until dissimilarity between each of the rehabilitated sites and the ARC (average reference community) reaches 0. The use of linear models limits the comparison of ARC-Dissimilarity results. As previously described (Chapter 2 Discussion, section 2.2), linear model coefficients are utilised here, despite their limitations, in order to prevent overinterpreting the relatively low number of samples for each rehabilitated site by visual overfitting of those samples. However, the limitations resulting from the imposition of linear models also carry over to the comparative assessment, preventing a more direct analysis of the shape of patterns in the data. This means that the assessment does not take into account whether the pattern of a rehabilitated site's convergence has been retained, only the similarity of the years to convergence estimates. This makes the assessment vulnerable to rating the agreement between a simplified analysis's estimate for a

rehabilitated site's years to convergence and that of the original Species Abundance analysis as high simply because it has a similar "time to convergence" when the actual path could be very different. From a management perspective, this is most problematic when sites display traits such as diverging post-convergence, such as DSC81 at Callide Mine. This is an important feature of DSC81's convergence pattern because it means it is now diverging and is potentially in need of management intervention. If trends such as this are not maintained across analyses, then important information is being lost even if the predicted convergence time does not drastically change. From a research perspective, this loss of information is more generally problematic, because if apparent trends in community composition differ across analyses then simplified analyses are severely limited in the kinds of studies they can be validly utilised in. The current ARC-Dissimilarity comparison does not test for such pattern retention, but future studies into the validity of these simplified analyses should take this important aspect into account.

As was previously alluded to in chapter 2, the use of linear models as part of the ARC-Dissimilarity analytical method is a reductive measure necessitated by the lack of data. In reality, we would expect rehabilitated site dissimilarity slopes to be asymptotic curves, as observed in a previous long-term monitoring study (Majer & Nichols, 1998), converging at first before levelling out at a dissimilarity to a reference site ARC similar to that of the reference site samples themselves, indicating the rehabilitated site has become fully alike to the reference site, including the various natural trends and oscillations that contribute to the reference site samples' average dissimilarity from the ARC, or the size of their "orbit" in community composition space. This proposed non-linear model approach would require much longer-running studies than the 4-6-year monitoring of Callide and German Creek mines; in the aforementioned study by Majer & Nichols (1998), the curves did not begin to plateau until 6-9 years into the study, and only one site appears to be plateauing out after 14 years of monitoring (Majer & Nichols, 1998). With the datasets we have, we must instead make use of linear models. The limitations of these linear models are made

particularly apparent by site TH91 at Callide Mine, which under the Large-Bodied Species Abundance analysis is diverging from reference site 6 but has a negative y-intercept, or negative dissimilarity. This is itself an impossibility and is not biologically meaningful, but even if we ignore any negative values, it suggests that site TH91 was identical to the reference site ARC some way into the rehabilitation program, before diverging.

We can acknowledge that the underlying assumption of the linear models, of convergence being equal to 0.00 dissimilarity from the ARC, is false and serves only as a useful approximation given the lack of data. But even when we do this, the models themselves remain unable to distinguish between sites that are not converging and sites that have reached the point of convergence and stabilised in terms of its dissimilarity to the ARC. This problem arises because nothing in the models relates position of the rehabilitated site samples to the position of the reference site samples. Site 8 at German Creek falls foul of this inability to distinguish between non-convergent and fully converged sites under the Large-Bodied scheme. Under the Large-Bodied scheme site 8 overlaps with reference site 3 in both NMDS and ARC-Dissimilarity but is simply counted as slowly divergent from reference site 3 in ARC-Dissimilarity, with the real significance of its positioning missed. This is a fundamental issue for the use of linear models, and also leads to the possibility of false equivalence between results in the comparison of ARC-Dissimilarity results. The solution, as previously discussed in chapter 2 (Discussion, section 2.2) and above, is to consider ARC-Dissimilarity and NMDS results in conjunction, and to evaluate the position of rehabilitated sites relative to the reference site samples as well as the ARC, either qualitatively or quantitatively, bringing vital context to the results.

While the use of linear models is a pragmatic choices which works within the confines of the limited datasets available, in an “ideal” dataset with sufficient data to fit the curved rehabilitation patterns we would expect to see if we had data for a full recovery, the current “time-to-convergence” method of evaluating ARC-Dissimilarity results for agreement would not be usable, as it is predicated

on linear models that either converge or diverge from a dissimilarity of 0. Instead, a more sophisticated model would be needed that compared results based on time taken until the slope of the curve becomes statistically indistinguishable from 0 (at a pre-defined p-value) and remains so, indicating the site has entered a stable orbit. However, given the amount of noise in ecological data, it would take a very long monitoring programme post-convergence to achieve a long-term trend that had a net dissimilarity change of 0, and that in itself assumes no other factors would cause the community composition to shift, independent of recovery, over such a long timescale, so this is a largely hypothetical problem concerning an ideal dataset. In terms of future studies, it would be better to consider linear-model-based "time to convergence with the ARC" and hypothetical curve-based "time until curve slope becomes fixed at 0" models as two ends of a continuum of modelling solutions, with longer-running datasets having to take account of curved recovery trajectories but adopting lower thresholds of convergence than the hypothetical model described above, such as when the curve approaches within a given distance to 0, or reaches 0 and remains within a given range around 0 beyond that time.

Ultimately, the ARC-Dissimilarity model used here, and the subsequent method used to compare results from this model, are limited options chosen due to the restricted dataset. They are imperfect models of the true recovery pattern, but the datasets they have been utilised on are themselves imperfect for capturing the true recovery pattern. As such, it is important for future studies to consider the extent to which their dataset captures the full extent of recovery and choose an appropriate variation of the models described and suggested here.

1.2 – Qualitative Interpretation of NMDS Ordinations

Given the value of NMDS ordination plots for providing information on rehabilitated site trajectories and the broader context within which they take place, the position of the other rehabilitated sites and reference sites, it is important to evaluate how the individual sites' trajectories hold up across the four simplified analyses. The simplest way to do this is qualitatively, testing

whether someone attempting to interpret the ordinations of each of the four simplified analyses would come away with the same conclusions as they would from the original Species Abundance analysis. However, this approach is subject to the risks of any qualitative assessment of data, particularly the subjective and potentially inconsistent interpretation of similarities between patterns, and should be treated with the requisite caution. To provide some fixed criteria of comparison between different analyses to limit this subjectivity, the focus of this exercise was whether sites were converging or diverging with the reference sites, the rate at which they were doing so and the consistency of that convergence/divergence, in addition to any key features of the patterns not included in that criteria, such as grouping with the reference sites (TH91) or on a clear intersecting trajectory rather than simply convergent (DSC81, TGC92), rather than judging whether patterns were the "same". This is similar to how capacity for subjective judgement was limited in the assessment of ARC-Dissimilarity results, but the qualitative interpretation has retained more scope for considering pattern similarity due to the more flexible criteria. Nonetheless, the focus on such criteria limits the evaluation of pattern similarity, and so is relatively forgiving of deviation in patterns between ordinations, provided the pattern is broadly similar in terms of direction – this leads to a lot of "erratic" patterns being lumped together under that label, as while they vary considerably in shape, the net result for convergence/divergence is often quite similar. A more flexibly qualitative approach would be better for distinguishing between these patterns, although such an approach would be even more difficult to replicate than the current approach. This is not necessarily particularly important for assessment of whether a site is converging or not, as the net convergence/divergence is the same, but end-users interested in the community patterns themselves rather than the net result may get very different results depending on the simplified analysis, and this has not been assessed. In future the evaluation could potentially be made more robust with repeated assessment by different evaluators, each scoring the different analyses according to how accurate they judged it, and those results then averaged, allowing for less

restricted qualitative assessments which take variation in pattern shape into account.

This comparative exercise, based on the 2D ordination plots, also highlights the problems that can arise when the axis explaining the most variation in the plot is not the primary axis of convergence (chapter 2 Discussion, section 2.1). At Callide Mine, the lack of any large-bodied ant species in the sample taken from TGB98 in the first year of the sampling program causes the ordination to be arranged around the difference between this sample and all other samples, meaning the reference and rehabilitated site samples overlap and we cannot tell if the rehabilitated site patterns in this ordination are a result of the Large-Bodied Species Abundance analysis, or simply that all other points are similar relative to TGB98 sample 1. In contrast, ARC-Dissimilarity is relatively robust to this sort of problem, as it only has 1 dimension, dissimilarity from the ARC, and unusually dissimilar samples will affect only the slope of the site that sample belongs to.

NMDS ordination is extremely useful to the assessment of mine-site restoration, but the inherently qualitative nature of its use in the context of assessment of individual rehabilitated sites' restoration patterns makes robust comparison across simplified analyses challenging.

1.3 – Inter-rater agreement: What makes a "good" simplified analysis?

In this study, inter-rater agreement analyses were used to assess how reliably the simplified analyses agreed with Species Abundance analysis on whether rehabilitated sites were converging or diverging and estimates for years to convergence. These analyses provide a score which is interpreted using classification systems taken from the literature. However, there are no standard reliability thresholds for either Cohen's Kappa or Intraclass Correlation Coefficients (ICC) (Bakeman *et al*, 1997, Koo & Li, 2016). The scores provided by these two analyses can be affected by a wide variety of factors. Kappa is used as a measure of rater (Analysis) accuracy but is also affected by the number of possible choices (in this case, 2, convergence or divergence) and the probability

of each choice (Bakeman *et al*, 1997). ICC scores are affected not only on agreement between raters but variability between the rated items, the sample size and the number of raters being tested (Koo & Li, 2016). Although Bakeman *et al* provide a formula for calculating the correct thresholds for Cohen's Kappa, it requires knowledge of the probabilities of convergence and divergence, which are unknown and likely to depend on complex ecological factors, and we could not find guidelines for our scenario for ICC (Majer, 1989, Bakeman *et al*, 1997). Instead, approximate rules of thumb were used. These are not a perfect fit for this study, but provide rough guidelines as to what counts as a "good" result. For Cohen's Kappa this could be supplemented by the calculation of the raw percentage agreement, to provide context to the Kappa scores. The range of appropriate ICC values for comparison of simplified analysis could potentially be calculated through modelling of simulated data with varying levels of predetermined agreement, as has previously been done for Orthopaedic research, to create a context-specific rule of thumb for good inter-rater agreement values, (Lee *et al*, 2012).

It is likely that the inter-rater reliability ratings for each of the four simplified analyses are lower than they would be based on inter-rater agreement alone, due to the lower number of samples, raters and possible choices. This is more of a problem for ICC ratings, where the score represents total agreement between Species Abundance analysis and each of the Simplified abundances, than for Kappa scores, as Kappa scores show reliable a simplified analysis is excluding agreement by chance, rather than all agreement, and can be compared to raw percentage agreement as a way of checking results (Cohen, 1960, McHugh, 2012, Koo & Li, 2016). However, it is still a substantial problem when using Kappa, as illustrated by the fact that, at Callide Mine, Kappa analysis (table 7) rates all four analyses as either moderately or poorly reliable, despite all four boasting raw agreement percentages in excess of 80%. This suggests a high portion of variation is being classed as agreement by random chance. In the context of finding a simplified analysis which has a high degree of agreement with Species Abundance analysis, a relatively low but significant Kappa score may still

correspond to a good degree of agreement, in the event that the expected random agreement is high.

The use of inter-rater agreement approaches does not provide a simple answer to questions of suitability or reliability of a simplified analysis, the interpretation of results being heavily context-dependent. Even once accurate, context-appropriate set of ratings have been established, depending on the use to which these analyses are put the acceptable rate of error may vary. Going forward, practitioners are encouraged to select a simplified analysis that meets the needs of their project, and the utility of these analyses is best considered in terms of conditions and contexts under which they work well and provide useful information, as opposed to a strict one-size-fits-all accuracy threshold.

2 – Discussion of Results

2.1 – Disagreement between comparison methods

There is considerable disagreement between the results of Procrustes analysis, Kappa analysis, ICC analysis and qualitative interpretation of NMDS ordinations regarding which simplified analyses were the most successful at replicating the results of the Species Abundance results. At Callide Mine, Species Presence/Absence analysis was highest-ranked based on NMDS Procrustes analysis (table 6) and ARC-Dissimilarity ICC analysis (table 8), but Functional Group Abundance analysis was the most effective in ARC-Dissimilarity convergence/divergence testing (Table 7) and Genus Abundance was most effective in NMDS interpretation (table 13), while at German Creek NMDS Procrustes (table 14) and ARC-Dissimilarity Kappa analysis (table 15) ranked Species Presence/Absence as most effective by a considerable margin, but none of the four simplified analyses were effective in NMDS interpretation (table 21) or ICC analysis (table 16). There were also broad consistencies – all three comparison methods found lower effectiveness of simplified analyses at German Creek than Callide Mine, Genus Abundance and Functional Group Abundance analyses scored similarly across all methods at both mines, except for ICC analysis of EYCA at Callide Mine, and the results of Species Presence/Absence

analysis were consistently found to be quite similar to those of Species Abundance. However, this broad consistency of results between the three assessment methods with variation in the particulars is not particularly surprising: the four methods are all based on the same inter-sample Bray-Curtis dissimilarity matrix, and Procrustes and NMDS Interpretation are both based on NMDS while Kappa and ICC are both based on ARC-Dissimilarity. But the four are used to assess different aspects of the community dissimilarity data.

The first major difference is between ARC-Dissimilarity and NMDS data. Kappa and ICC analyses of ARC-Dissimilarity data on the actual dissimilarities between site samples, and only looks at the dissimilarities between site samples and the ARCs of the reference sites. No other information goes into the analyses. In contrast, Procrustes and NMDS interpretation are built on NMDS ordination, which does not directly map onto the dissimilarities between sites, instead being based on rank-ordered dissimilarities between samples, and furthermore taking into account dissimilarities between *all* samples and attempting to convey those relationships as accurately as possible in two dimensions, when a fully accurate display would likely take many more. This makes the NMDS ordination several steps removed from the dissimilarities it is derived from, and we should not expect the two to map onto each other 1:1, and hence should not be too surprised when results differ between methods, particularly when the differences between simplified analyses are not always especially large at each mine.

Differences in ratings between Kappa analysis of inter-Analysis agreement on site convergence/divergence and ICC analysis of agreement convergence time are likely to be due to differences in the required level of precision for high agreement in each analysis. At the convergence/divergence level, all that is required to achieve high agreement is agreement on the direction of community change the site – getting more or less similar. In contrast, high agreement for actual convergence time requires far more precise agreement on how quickly the site is changing and how dissimilar the site originally was from the reference site

ARC, so it is unsurprising that ICC results are frequently so much lower than Kappa results at both sites.

As for Procrustes and NMDS interpretation, the difference probably lies in the fact that Procrustes analysis takes into account the entire ordination, while NMDS interpretation focuses on individual site trends. This means Procrustes analysis also takes into account the overall positions of sites, such as the sorting of reference and rehabilitated sites on one axis, which is fairly consistent across all analyses at both mines. NMDS interpretation on the other hand focuses only on the fate of individual sites, where relatively small changes in position of individual points within a cluster are sufficient to change the story of that site between analyses.

Differences in results between assessment methods are not that surprising, especially as it may take relatively small changes in fidelity to cause changes in ranking of the four simplified analyses in each of the four analyses. It is best to consider the four methods as complementary approaches, revealing different aspects of the four simplified analyses and how they alter the story coming out of the data, and, when selecting a simplified analysis, to prioritise the results of assessment methods that tests the aspects of the simplified analyses that are of interest.

2.2 – Behaviour of simplified analyses

The four different simplified analyses differed from Species Abundance in various ways in terms of their underlying mechanics, and this is reflected in consistent differences, trends and tendencies in the of the four from Species Abundance analysis and from each other across mines.

Both Genus and Functional Group Abundance analyses had very similar ARC-Dissimilarity patterns and NMDS ordinations at both mines (tables 6, 7 & 8) They both consistently had lower Bray-Curtis dissimilarities between rehabilitated sites and the ARCs of the reference sites than Species Abundance analysis both

at the beginning and end of the sampling period, and less clear separation of reference and rehabilitated sites in NMDS ordination (figs. 5 & 6). They also both placed the three reference sites much more closely together in NMDS ordination than Species Abundance (figs. 5 & 6). This consistent reduction in dissimilarity between sites compared to Species Abundance is probably a product of these two simplified analyses being based on higher taxonomic or ecological ranks than Species Abundance, grouping species together into genera and functional groups respectively. This makes these analyses less discerning than Species Abundance as sites are more likely to share genera or functional groups than individual species, simply because there are multiple species that count towards a genera being present at a site and while certain species may be specialists on highly disturbed rehabilitated sites, mid-recovery rehabilitated sites or fully restored/reference sites, higher groupings are more likely to collectively cover a greater range of conditions and have representatives at multiple stages in the recovery process, either due to having multiple specialists on different points the recovery gradient or due to a higher chance of having a generalist species within their ranks. This makes these higher rank analyses less able to discern between sites so inter-site dissimilarities are lower.

That reduced dissimilarity also applies to reference sites, which are even more likely to share broad patterns of environmental conditions, and subsequently community composition, and be separated by the fine-grain detail of Species Abundances rather than higher ranked groups, than reference and rehabilitate sites are. This means that under these two analyses the reference sites are clustered together, and consequently convergence patterns on the three reference sites tend to be more homogeneous than under Species Abundance analysis. At Callide Mine, the relative position of the three reference sites compared to the rehabilitated sites is most similar to the position of reference site 6 under the original Species Abundance ordination, and the patterns for the three reference sites under Genus and Functional Group Abundance analyses subsequently most closely resemble those of reference site 6 under Species Abundance (fig. 1 (Chapter 2 Results), figs. 9 & 10). But the reference sites need

not necessarily cluster on the original position of one – at German Creek the three reference sites still cluster together, but cluster in the centre of the ordination rather than on a position any of the three held under Species Abundance analysis.

The general behaviour of Large-Bodied Species Abundance analysis is harder to identify, due to the problems associated with the Callide Mine Large-Bodied Species Abundance Ordination. The increased tendency of disruption to the NMDS ordination (fig. 7) is in itself an important characteristic of Large-Bodied Species Abundance analysis and is discussed in detail in chapter 4 (section 1.2). Beyond that it is hard to make any particular inferences about its behaviour beyond the fact that it is quite variable. Despite having a high level of agreement in ARC-Dissimilarity convergence/divergence results at Callide Mine (, fig. 11), it achieved less than 50% agreement at German Creek (, fig. 15), which, together with the issues arising at Callide Mine, might suggest that it is the simplified analysis most vulnerable to idiosyncratic effects, perhaps as a result of its unusual simplification method, which excludes a variable portion of the dataset based on a single characteristic which was selected due to its practical convenience rather than an ecological basis and so does not guarantee the set of species selected as the basis for analysis are important in or representative of site community dynamics (Andersen, Hoffmann, & Sommes, 2002).

Species Presence/Absence was the simplified analysis that most closely and consistently resembled Species Abundance in its results, both in ARC-Dissimilarity (tables 7, 8, 15 & 16) and NMDS ordination (tables 6 & 14), consistently having the lowest Procrustes score and the only simplified analysis to have a consistently high percentage agreement for convergence/divergence classification. However, Species Presence/Absence appeared to be more conservative than Species Abundance in how dissimilar sites were from each other, with reference and rehabilitated sites more clearly separated and the reference sites clearly separated from each other in NMDS (figs. 8 & 16). The rehabilitated sites appeared to be slightly more clustered with each other

however. Species Presence/Absence places increased priority on which species are present or absent at sites and less on the abundances of species. This means there is an increased emphasis on the percentage of species shared between sites and a reduced emphasis on the abundances of shared species. The likely consequences of this is greater differences between groups of sites which are less likely to share species (e.g. rehabilitated vs reference sites) and greater similarity within those groups, hence we see greater clustering between rehabilitated sites and between reference sites (and site TH91 at Callide Mine), but and greater separation between reference and rehabilitated sites at both mines. The clustering of reference sites is nowhere near as extreme as under Genus or Functional Group Abundance however, as Species Presence/Absence merely emphasises similarities between them which already exist in the Species Abundance data, rather than increasing those similarities with broader groupings. The fact that Species Presence/Absence was the most consistently in agreement with Species Abundance, and uniquely achieved a high level of agreement for estimated years to convergence at Callide Mine, suggests that the presence or absence of species across sites is an important component of species abundance data, and hence patterns are not drastically altered by the loss of abundance data for each species. These results also demonstrate that, although most simplified analyses were unable to meet the required level of precision for high estimated years to convergence agreement, it is at least possible to do so under the right conditions, meaning estimated years to convergence is, at least theoretically, an achievable level of agreement for simplified analyses, even if it is rare in practice.

2.3 – Differences between sites

One of the most striking features of the results is the stark difference in performance of all four simplified analyses between the two mines. At Callide Mine all four analyses appear to be fairly faithful replicas of the Species Abundance analysis in terms of convergence/divergence results (table 7), and in NMDS results as well (table 6 & 13), Large-Bodied Species Abundance analysis notwithstanding owing to issues discussed in section 1.2. But at German Creek

only Species Presence/Absence analysis emerged as being highly accurate to the Species Abundance results (tables 14 & 15) or even accurate more than 2/3s of the time, but performed as poorly as the other three in ARC-Dissimilarity estimates of years to convergence and NMDS interpretation (tables 16 & 21). These results are important, because they suggest that none of the four simplified analyses are necessarily applicable in all contexts but that, by the same token, they are all applicable in at least some, and identifying the conditions for their successful usage is of critical importance. There are a number of possible explanations for this contextual usefulness.

The first possibility is essentially the null hypothesis for the differences between mines – that the results for the two mines are not actually significantly different. The number of rehabilitated sites surveyed at both sites is relatively low – 6 at German Creek and 8 at Callide Mine. If we consider each site as a test in the experiment, then those are sample sizes of 6 and 8. At Callide Mine, the average fidelities of each simplified analysis for estimated years to convergence was approximately 3/8 sites (2.67-3.33) (table 8), and for convergence/divergence the averages were approximately 7/8 (6.67-7) (table 7). At German Creek the corresponding results are 1-2/6 (1.33-2.33) (table 16) and 3-5/6 (2.66-4.66) (table 15). While in terms of percentage accuracy they are quite different, in practice they are not huge differences in numbers of sites. Having said that, in ARC-Dissimilarity convergence/divergence testing all four analyses do worse at German Creek (tables 7 & 8, 15 & 16), and all but Large-Bodied analysis do worse in Procrustes (tables 6 & 14) and NMDS interpretation (tables 13 & 21) (though in ARC-Dissimilarity estimated years to convergence testing, Large-Bodied analysis does equally well, and Presence actually does slightly better (table 16)). The large differences in agreement of Genus and Functional Group Abundance analyses and Species Presence/Absence in Procrustes and NMDS Interpretation also suggest there is a significant difference in performance between the two mines. It is therefore unlikely, but possible, that the difference between the sites is simply a matter of coincidence. This hypothesis can be easily tested by subjecting more mine-site rehabilitation datasets to these simplified analyses.

A more interesting, although still site-specific, possibility is that the trends at Callide Mine are simply easier to detect. The four analyses all "simplify" the data collection process by reducing the amount of information. Genus Abundance and Functional Group Abundance analyses reduce the resolution of the data, shedding the species-level information and with that information species-level trends. Species Presence/Absence analysis cuts out the abundance aspect of species-level data, and Large-Bodied Species Abundance analysis uses the abundance data of only a minority of the species present. As such, all four are, to varying extents, less sensitive to changes in ant community composition than Species Abundance analysis. While each may differ in the size and type of changes in community composition each is insensitive to, all four are less sensitive. Consequently, smaller changes in community composition, particularly those based on abundances of species alone, will not register in the data collected for some or all of these simplified analyses. And, comparing results of ARC-Dissimilarity and NMDS Ordinations between mines, changes in community composition at Callide Mine are much bigger than at German Creek. Rehabilitated sites undergo large, rapid changes in community composition over the course of the study which alter the functional group profiles of the rehabilitated sites, and are clearly converging or diverging with the reference sites. In contrast, changes occurring at German Creek are much less dramatic, without major changes to the functional group profiles of the rehabilitated sites and so are much less likely to register in the simplified analyses. This hypothesis could be tested by comparing success rates of the simplified analyses between more or less successful rehabilitation projects, and by re-examining the success rates of the simplified analyses at replicating trends of particular rehabilitated sites at Callide and German Creek and cross-referencing those success rates with which sites have the biggest changes in dissimilarity from the ARCs, or the largest movements on the NMDS ordinations, in the original Species Abundance analysis.

Of course, under particular circumstances certain analyses may still be particularly sensitive to changes, and this is the third hypothesis – that differences between results of the simplified analyses at each site are driven by ecological differences between the mines' ant communities. The specific ecological factors underpinning the effectiveness of each simplified analysis will be discussed in more detail in the general discussion, but the broad theory will be outlined here. Different locations differ in their environmental conditions, and hence in their species occurrence and community compositions, even if those differences can sometimes be only small (Andersen, 1995b). This means that that the ant community composition of different sites may be differentially organised, such that different amounts of information are lost by the transition to one of the simplified analysis at each site, and consequently the success of each analysis differs between sites as their correlation to the original analysis varies with local ecological conditions. In opposition to this idea, all four analyses experience a reduction in performance at German Creek relative to Callide Mine in at least one analysis, suggesting a common cause. It is however possible that the circumstances at Callide Mine are such that the community composition organising conditions happen to favour all four analyses, and certainly it can be argued from the results that Large-Bodied analysis did not really perform too differently at either mine and the three other analyses all behaved similarly to each other at both mines, suggesting a common driving factor in their patterns after all. But the failure of Large-Bodied analysis at Callide Mine can largely be attributed to incompatibilities in the procedure and assessment methods regarding site samples where target species are absent, rather than ecological causes. When these procedural issues were avoided, by using ARC-Dissimilarity assessment, Large-Bodied species performed similarly to the other analyses and had the greatest difference in performance between the two mines of any of the simplified analyses in ARC-Dissimilarity convergence/divergence testing. This hypothesis can be tested by assessing whether the qualities of the composition of the ant community at the Callide mine sites favour the four simplified analyses. The qualities that favour improved fidelity in the four analyses will be outlined in the general discussion (section 1).

Both alternative hypotheses rest on the two sites differing in on-site ecological conditions, either in ways that cause the recovery at Callide Mine to proceed more rapidly, creating a stronger and hence more easily replicated signal, or in ways that the community is structured such that what changes do occur are less likely to be detected by simplified analysis procedures. The most likely driver of such differences is the climate of each site. Differing climates will as a matter of course shape and affect the local ecosystem in various ways, creating different conditions and niches as a consequence of changes in variables such as temperature and water availability, as well as secondary effects of differences in those variables such as variation in vegetation structure and net primary productivity, all of which have been linked to ant community structuring in Australia and elsewhere (Majer, 1989, Majer, 1992, Andersen, 1995, Majer, 1996, Kaspari, O'Donnel & Kercher, 2000, Dunn *et al*, 2009, Kwon & Lee, 2015). Temperature in particular has been identified as a fundamental organising factor for ant activity and distribution of ant species and in relative abundances of functional groups and so is a likely candidate for driving differences in community structure between the two sites (Andersen, 1995, Dunn *et al*, 2009). However, while rainfall is less important than temperature in the general structuring of ant communities, in the context of rehabilitation it becomes a key climatic variable, defining the rate of return of ant species to rehabilitated mine sites (Majer, 1989, Majer, 1992, Majer, 1996, Kaspari, O'Donnel & Kercher, 2000, Dunn *et al*, 2009). Higher annual rainfall drives faster increases in species richness at rehabilitated sites and so consequently should drive faster convergence of rehabilitated site communities with local natural communities observed at reference sites (Majer, 1989, Majer, 1992, Majer, 1996). This pattern is particularly well-established in Australia, with a clear continuum of increasing rate-of-return observed from the relatively dry Mediterranean biogeographic region in the south to the very humid Tropical Monsoonal regions in the North (Majer, 1989). As such, it follows that if there is a difference in rainfall between German Creek Mine and Callide Mine, it could be driving the greater recovery rates at Callide Mine. However, Callide and German Creek Mines are only

approximately 270km apart and have very similar climates, at least in terms of annual rainfall (Callide = 684mm, German Creek[3] = 663m) and mean annual temperature (Callide[4] = 29.1°C, German Creek[4] = 29.3°C), so these key climatic factors are unlikely to be causing substantial differences between the ant communities of the two mines, whether that be differences in the community structure or in rates of recovery, and so we must look to other factors to explain these results (Bureau of Meteorology, 2014).

These are three potential hypotheses to explain the differences in simplified analysis performance across the two mines, and the two alternative hypotheses are not mutually exclusive, as the difference in simplified analysis fidelity between the two sites could be driven by both a lack of major changes at German Creek (hypothesis 2) and a community composition and community compositional changes that made detection of changes that did occur difficult for the simplified analyses (hypothesis 3). In fact, an examination of the data suggests that both causes are likely to be in effect. One of the key differences between convergence patterns under Species Abundance analysis and those under the simplified analyses at German Creek is the lack of the clustering of rehabilitated sites 9, 10, 11 & 12 and their shared trajectories under the simplified analyses (Chapter 3 Results, section 1.7). As discussed in chapter 2 (Discussion, section 3), these patterns are driven by the alternating dominance of these rehabilitated sites by two species of Iridomyrmex (rufoniger group), I. reference site 3 between years as they alternate between sharing dominance by species B & I. species E, and their subsequent large changes in similarity to *I*. species *E* with the reference site and being instead dominated by *I*. species *B*.

[3] German Creek annual rainfall data taken from nearest weather station for which annual rainfall data was available, at Boroondara, station number 035109, 12.6km away

[4] Callide and German Creek annual mean temperature data taken from nearest weather stations for which annual mean temperature data was available. Callide data taken from Thangool Airport, station number 039089, 19.3km away. German Creek data taken from Clermont Post Office, station number 035019, 96km away

However, these changes, which play a key role in shaping changes in community dissimilarity under Species Abundance analysis, are entirely driven by abundances of two very closely related species. This means that their impact on the results of Species Abundance analysis is completely undetected in all four simplified analyses. Under Genus Abundance analysis, they are both members of the genus Iridomyrmex and so the switching between the two species is not detectable, and this also applies at the Functional Group Abundance level. Under Species Presence/Absence, their hyperabundance is not accounted for, and so the switch between the two species is simply two species of many swapping over or even co-occurring, as in a number of samples at rehabilitated site 12 and 11 both species are present simultaneously, just with one at high abundance and one at low abundance. Finally, neither species is a large-bodied species, and so they have no effect on Large-Bodied Species Abundance analysis. As such, an important component of the convergence patterns of Species Abundance at German Creek is completely excluded from all four simplified analyses, and this likely explains why all of the simplified analyses perform less effectively at German Creek, where much of the pattern is define by changes in species abundances alone. This is a particularly extreme example, but the loss of Species Abundance data also affects other aspects of the analysis of German Creek. At reference site 7 at German Creek, there is a revolving community of Iridomyrmex species (see chapter 2, section 2.2) which rapidly displace each other. This means that, at the Species Abundance level, reference site 7's ARC has a very low median Iridomyrmex abundance comprised of those species that maintain a presence throughout the sampling period rather than a high one reflecting the continuous high abundance of at least one Iridomyrmex species throughout. Under Genus Abundance & Functional Group Abundance analyses, that changes, with all Iridomyrmex species being lumped into the genus-wide count or the Dominant Dolichoderine count respectively, so that the key factor is the abundance of the collective rather than of individual species. This means these two analyses are more robust than Species Abundance analysis to similar species which competitively displace each other during recovery of rehabilitated sites (Majer, 1989). However, it will also alter the outcome of dissimilarity calculations

between the ARC and other samples, as the abundances of the Iridomyrmex species now have a much bigger contribution to the ARC composition. The decreased agreement at German Creek compared to Callide Mine is not necessarily due entirely to these two examples, but the relatively static functional group profiles of the sites at German Creek indicate that within-group species abundance changes and displacement by similar species, as within *Iridomyrmex*, are likely to be important more generally, and this likely explains the reduced performance of simplified analyses at this site and the relative success of Species Presence/Absence analysis. If changes in species composition are the driving force for changes, Species Presence/Absence would better reflect changes in the community at the Species Abundance level, though it would still be insensitive to community changes driven by changes in abundances of species as described here.

Hence the likely cause of the reduced performance of the four simplified analyses at German Creek relative to Callide Mine is a combination of both alternative hypotheses – that the community changes occurring during recovery at German Creek are ecologically "small" changes in community composition consisting of changes in species abundances, rather than changes in functional group abundances representing major ecological shifts, particularly the competitive displacement of closely related *Iridomyrmex* species, which are not detectable by the four simplified analyses for various reasons. This means that the successful use of simplified analyses is context-dependent, requiring an understanding of the local ant community and its dynamics and whether a particular simplified analysis is reflective of those dynamics.

Chapter 4 – General Discussion

1 – Evaluating the use of simplified analyses

The results of this study suggest that the successful use of simplified analyses of ant communities for assessment of mine-site rehabilitation is highly context-dependent. While all four analyses have shown some potential to be successful substitutes for a full, Species-Abundance-based assessment, all four have also been shown to be inconsistent in how closely they map to a full assessment, and consequently their potential utility is severely limited due to their unreliability. However, that is not to say that these analyses cannot be made useful. Instead, their use is context-dependent, suited only to certain ecological and management contexts where they can be used effectively.

1.1 – Genus Abundance and Functional Group Abundances

Genus Abundance and Functional Group Abundance analyses showed extremely similar results across both mines and all four assessment methods at each mine, suggesting the results of both analyses are driven by the same factors. This is unsurprising, as the nine functional groups are made up of genera, and the classification of genera into one of those functional groups is in part dependent on each genus's habitat preferences at the biogeographic scale (Andersen & Hoffmann, 2003a). Disturbance involved in mine-site creation and subsequent rehabilitation and recovery is the disturbance most closely resembling biogeographic-scale environmental differences, owing to the extreme disturbance and succession-like changes in the ecosystem that occur during the recovery process, with subsequently similar changes in ant functional group composition during the succession (Andersen, 1993, Bisevac & Majer, 1999, Andersen & Hoffmann, 2003a). This means that, at mine-sites, the functional groups are effectively groupings of genera with similar responses to the environmental changes occurring during recovery, so we should expect very similar patterns.

Functional Group Abundance does not differentiate between reference and rehabilitated sites as well as Genus (figs. 5 & 9, 6 & 10), and has more sites overlap in dissimilarity from the ARC with the reference site samples. This is unsurprising as the reduction in information, summarising the trends of many genera into only nine functional groups, would result in a loss of power to differentiate between the sites. Having said that, both of these simplified analyses grouped the reference sites far closer together than in Species Abundance, with a subsequent homogenising effect on ARC-Dissimilarity results across reference sites, so the improvements offered by Genus Abundance analysis over Functional Group Abundance analysis in this aspect are marginal.

What elements of community structuring lend themselves to accurate use of Genus Abundance and Functional Group Abundance analysis? As with the Genus Abundance ~ Functional Group Abundance relationship, the answer is likely to lie in consistency of responses among the composite species and species-groups of the genera at a site being surveyed. If the species within a given genera all have similar ecological requirements and responses to disturbance, then it follows that the agreement between the genus-level changes in presence and abundance and those at the species level will be high. Therefore, the higher the average homogeneity of species within genera at a site in terms of ecological requirements and disturbance responses, the higher the congruence of results between species-level and genus-level analyses will be, and subsequently between species-level and Functional Group-level analyses. However, there are additional complexities. The abundance of ant species that are common or occur in large numbers, such as members of the *Iridomyrmex rufoniger* species group at Callide & German Creek mines, has a greater influence on the results of abundance-based analyses than those of rarer or solitary foraging species. As such, the homogeneity of species responses within genera of abundant ants such as *Iridomyrmex* will be more important than the homogeneity of species responses in genera of less abundant ants, and likewise the homogeneity of responses among abundant members of a genus are more important than the homogeneity of responses of less abundant members. In the case of *Iridomyrmex*

in particular, the problem is simplified, at least at German Creek Mine and Callide Mine, by a single hyperabundant species frequently accounting for the majority of *Iridomyrmex* abundance at a site, and frequently far higher percentages, resulting in high response homogeneity. Unfortunately, the practical application of this hypothesis is limited, as it means the utilisation of Genus Abundance or Functional Group Abundance analysis is dependent on knowing 1) all the species present at the site, and 2) their degree of homogeneity in habitat needs and disturbance response. While the first condition could probably be met by an initial species survey at the start of the monitoring programme, the second condition requires an in-depth understanding of the ecology of the species present at the site. This not only puts these two simplified analyses beyond the use of non-specialists in ant ecology, a major flaw given one of the great advantages of these approaches is the fact that genera can be relatively easily identified without specialist knowledge (Andersen, 1990), but for the hyper-diverse and relatively poorly understood Australian ant fauna, such species-specific data is, by and large, non-existent, assuming species at the site have even been formally identified (Andersen & Hoffmann, 2003a). However, for studies in less speciose and better-studied areas such as Europe, assessment of projects for suitability of genus- and Functional Group-level analysis is a possibility (Ottonetti, Tucci & Santini, 2006), and while the necessary inclusion of specialists at the early stages of a project to carry out such assessment is a limitation, it is still less limiting than the requirement of a specialist's involvement in the entire monitoring programme, as is required under a full species-abundance assessment.

This ecological context-dependency is likely why there is so much contention over the use of genus abundance analysis in the literature (Souza *et al*, 2018). While many authors, including Souza *et al*, report Genus Abundance analysis to be effective at reproducing the information conveyed by species-abundance analyses (Pik, Oliver & Beattie, 1999, Nakamura *et al*, 2007, Ribas and Padial, 2015), others have found otherwise, finding genus to be a variable and unreliable indicator of species-level trends, particularly for ants (Andersen, 1995, Rosser &

Eggleton, 2012). In particular, it has been found that the effectiveness of genus as a surrogate measure is conditional on low species:genus ratios, and on study site, even at quite small scales, a finding supported by this study (Andersen, 1995b, Lovell *et al*, 2007, Grantham *et al*, 2010, Ribas and Padial, 2015). The improved accuracy of Genus Abundance analysis at low species:genus ratios is in accordance with my prediction that greater homogeneity of species responses within the genus will lead to improved performance of this analysis, as variability between species in a genus with only a small number of representatives is likely to be lower so congruence between the two levels of analysis will be higher (Lovell *et al*, 2007).

In contrast, the use of the functional group scheme has generally been found to be an effective simplified analysis, particularly for the assessment of mine-site recovery (Andersen, 1993, Bisevac & Majer, 1999, Pik, Oliver & Beattie, 1999). However, when it is compared to other simplified analyses, it has generally found to be less faithful to species abundance results than analysis at the genus level (Pik, Oliver & Beattie, 1999, Nakamura *et al*, 2007). Functional group analysis of disturbance is often treated as somewhat distinct from species-level analysis, with the ratio of the different functional groups at each site being compared across sites or site types, so evaluation of the factors affecting its performance as a disturbance analysis have generally been limited to assessment of factors that affect whether functional group analysis can detect disturbance on its own merits, rather than in terms of its relationship with species abundance (Andersen & Hoffmann, 2003a, Andersen, 2017). This study helps to bridge that gap, the extremely close linkage of genus and functional group analysis suggesting its correlation to species abundance analysis is dependent on the same factors as Genus Abundance analysis, and furthermore that Functional Group Abundance patterns of change can potentially correspond quite closely to species-level patterns, although, as with genus, this is heavily context-dependent.

Taking all of this into account, under what management situations might we practically be able to make use of genus or functional group analyses? As

previously mentioned, the use of Genus Abundance and Functional Group Abundance analyses is limited by the details of the species-genus relationships, and that in ants that relationship is highly variable in genera and space (Andersen, 1995b, Andersen & Hoffmann, 2003a, Rosser & Eggleton, 2012). Hence their use is predicated on a solid understanding of the ecology of species and genera, something that seriously limits their use in hyper-diverse regions such as Australia, but makes them potentially quite viable in less speciose regions. A potential mitigating factor is that low species-genus ratios reduce potential for variability in species response within genera, and so this comparatively easily assessed aspect of community structure may potentially be useable as a proxy of species-genus ecological response correlation, in order to assess the suitability of Genus Abundance and Functional Group Abundance analyses for a given site and project. While the main potential utility of Genus Abundance and Functional Group Abundance analyses is whether they can be used in place of Species Abundance assessment, they also have a more niche use. A major issue with both NMDS and ARC-Dissimilarity approaches to rehabilitation assessment is the inability to verify whether a site is not recovering, or is simply not converging on the reference sites selected for the study. Under a species-abundance assessment, the number of reference sites is understandably limited by time and cost, and a representative set must be selected prior to the study as likely candidates (Andersen & Hoffmann, 1997, Andersen & Hoffmann, 2001b). However, in the event that many sites do not appear to be converging on the selected reference sites (e.g. at German Creek), these higher-taxon approaches can potentially be utilised as a means to rapidly assess a wider variety alternative local ecosystems to identify potential alternative reference sites, as while they are not necessarily sufficiently accurate for long-term monitoring of trends, they are likely sufficiently accurate to detect if convergence is occurring. Or, if monitoring begins after colonisation and community development has already begun on at least some rehabilitated sites, such as at Callide and German Creek mines, then a rapid Genus Abundance- or Functional Group-level "snapshot" assessment prior to beginning the full monitoring programme could be effective in highlighting which reference sites

rehabilitated sites are converging on, if any. However, even these uses are subject to community composition, and it is unlikely that Genus or Functional Group Abundance could ever be utilised without at least some previous Species Abundance surveying of the area to determine suitability. Their relative ease of use makes these two simplified analyses well-suited to interim surveys of conditions, looking for broad trends, even if the circumstances mean they are not suited for use in many areas and the identification of such areas is challenging.

1.2 – Large-Bodied Species Abundance

Large-Bodied Species Abundance analysis was the weakest of the four analyses overall, consistently performing low in Procrustes analysis and NMDS interpretation and performing worst of the four analyses in ARC-Dissimilarity convergence/divergence testing at German Creek. However, Large-Bodied Species Abundance analysis still successfully replicated the ARC-Dissimilarity convergence/divergence trends of the Species Abundance analysis at Callide Mine 83% of the time. This suggests the use of Large-Bodied species is a highly conditional one, but one that, like the other three simplified analyses, has the potential to be viable.

Mechanically, Large-Bodied Species Abundance analysis differs from the other three analyses in one crucial aspect. Unlike the others, Large-Bodied Species Abundance analysis does not make use of the full ant community dataset on at least some level. Genus Abundance and Functional Group Abundance analyses, although they merge the abundances of the species into higher taxonomic units, still make use of all those abundances, and Species Presence/Absence analysis still makes use of every species even if it does not take into account their abundances. This means that Large-Bodied Species Abundance analysis is in some ways the most different from the Species Abundance analysis in terms of data input. On the other hand, it is the only analysis to still use full-resolution species abundance data, meaning it retains the greatest taxonomic precision of the four analyses. In practice, it would appear that the distancing effect of losing

most of the data outweighs this unique capacity for matching precision in this study. Restricted Species Checklists such as Large-Bodied Species Abundance analysis are limited by a few key factors, the first and most important being extent of information loss (Grandin, Lenoir & Glimskär, 2013). At Callide and German Creek Mine, the use of the Large-Bodied species criterion resulted in reduction of species in the dataset by 66% and 73% respectively. Vellend, Lilley & Starzomski (2008) found random removal of 10% of species in a list reduced correlation between restricted-species and all-species community data in a variety of taxa by no more than 0.2, but reduction by 50% led to reductions ranging from 0.3 to 0.9, with invertebrate taxa including ants tending towards larger reductions. Grandin, Lenoir & Glimskär (2013) similarly found that reductions of 37% and 84% of plant species in Swedish grasslands were sufficient to produce quite different ordination patterns, and concluded the size of the restricted species list was more important than the species selection criteria. At German Creek, the loss of key species contributing to the dissimilarity between sites (*Iridomyrmex* sp. *E* & *B*, *rufoniger* group) due to their small size was the major factor. At Callide Mine, where the patterns of community change were not so heavily influenced by two species excluded from the restricted species list of species above 4mm, the average convergence/divergence ARC-Dissimilarity matching fidelity at Callide Mine are twice that of German Creek, and are similar to those achieved by the other simplified analyses. Furthermore, previous research on the use of the Large-Bodied Species Abundance analysis on the diverse Australian ant fauna was able to successfully replicate results of a full species abundance survey under comparable conditions of loss of 76% of species and with even less information, as only presence/absence data was used rather than abundance (Andersen *et al*, 2002), although it is worth noting that the results replicated by Andersen et al are much more general than the results we were attempting to replicate here, which are actually quite precise relative to other studies on simplified.

This suggests that, like Genus and Functional Group Abundance analyses, the use of Large-Bodied Species Abundance is dependent on ensuring the main drivers of

community change patterns are ones that are detectable by this simplified analysis. More generally, the question of how representative the species composition of the restricted checklist of species used in Large-Bodied Species Abundance analysis is an important one for the effectiveness of this analysis – if the species list is not representative, either because, like at German Creek, it excludes key drivers, or due to a more diffuse lack of representation of the wider ant community at a site in terms of habitat preferences and restoration response, then this analysis will be ineffective as a monitoring tool. Conversely, if the species on the list were the species at the site that were most sensitive to the changes wrought by restoration and so were the main explanatory variable in the community composition change patterns displayed, then we would expect close correlation of results between Species Abundance analysis and Large-Bodied Species Abundance analysis. Hence Large-Bodied Species Abundance, and more broadly restricted species checklists, have the most variable potential agreement reliability, potentially having extremely high concordance with Species Abundance if the right species are selected, or none at all if the wrong ones are. In the case of the species selected for the Large-Bodied Species Abundance list, having been selected on the basis of a functional morphological trait, there is likely to be at least some clustering of species within this group in terms of response to environmental factors, both as a result of phylogeny and ecology (Crisp *et al*, 2009), so their effectiveness as a representative sample of the ant community is questionable, unless the disturbance they are being used to monitor is one that large-bodied species are particularly sensitive to.

Setting aside the inter-site variation of this study, Large-bodied species analysis also suffers from a few additional across-the-board issues as a result of its uniquely limited nature. Firstly, as demonstrated at Callide Mine and previously discussed (Chapter 2 Discussion, section 2.1), the use of a restricted species list means running the risk of no species meeting the list's criteria, resulting in a site with "0" species (Clarke, Somerfield & Chapman, 2006). A site having no species that meet a certain criterion is very different thing ecologically than a site with no species within the taxa of interest, and yet within ecological analytical tools

such as Bray-Curtis they are treated identically unless specifically corrected for, as in this study. Even when corrected for, they can still cause severe disruption to important analytical methods in the field such as NMDS ordination, to the point of rendering the ordination useless for interpretation (fig. 7 & Chapter 3 Results, section 1.4). This issue is compounded by the fact that mining and rehabilitation are some of the strongest disturbance types for epigaeic ants, due to directly disturbing the soil they nest in, as opposed to indirectly affecting them through vegetation removal. Thus, monitoring immediately after rehabilitation is likely to find few or no species even without using a restricted checklist, increasing the likelihood of such a scenario arising (Andersen & Hoffmann, 2003a, Clarke, Somerfield & Chapman, 2006). The focus on Large-Bodied ants in particular is also problematic on a sampling level, as large ants, as a rule are less likely than smaller species to be caught in pitfall traps (Marsh, 1984, Olsen, 1991).

All in all, Large-Bodied Species Abundance analysis can be problematic to use, as a result of issues arising from being a restricted species checklist and from the selection criteria used for that list. It involves the exclusion of a large proportion of the species in the dataset, and with them much of the data, a reduction which has been shown in this study and elsewhere to result in very different community composition patterns when compared to a full species abundance analysis. Due to the selection of species on the basis of a morphological character that appears to be phylogenetically clustered, all the species in the list are likely to share at least some degree of concordance in habitat requirements and response to environmental changes (Crisp *et al*, 2009), which also reduces the likelihood that they will have a range of responses representative of the whole community. On top of these theoretical issues, the focus on Large-Bodied species involves a number of practical issues, including the increased likelihood of generating "false negatives" in community data and the consequent analytical problems that causes, compounded by the fact that as a group these species are less likely to be collected through pitfall trapping, the most common and simplest sampling procedure for ants (Bestelmeyer *et al*, 2000, Andersen & Hoffmann, 2003a). As such, while previous work has shown Large-Bodied Species

Abundance analysis can be successfully utilised and can drastically reduce processing time (Andersen *et al*, 2002, Arcoverde *et al*, 2017), and has shown similar success at Callide in replicating convergence/divergence results, at the current time there is no reason to recommend it over the more consistently faithful, more user-friendly and less specialist-dependent Genus Abundance and Functional Group Abundance analyses for general mine-site rehabilitation and restoration assessment. However, there is an exception to this recommendation – due to their shared functional trait, the group as a whole are particularly sensitive to disturbance (Ness *et al*, 2004, Gibb *et al*, 2018). Previous work with Large-Bodied species has shown them to be sensitive to changes in SO_2 where functional group composition was not, and so Large-Bodied Species Abundance analysis may have a role as a specialist simplified analysis for such cases, pending further investigation (Andersen *et al*, 2002).

1.3 – Species Presence/Absence

Species Presence/Absence analysis showed the most potential of the four simplified analyses as a general simplified assessment measure, consistently ranked as the most similar analysis in Procrustes analysis, and was the only one of the four to maintain a high degree of fidelity to the original Species Abundance analysis in ARC-Dissimilarity convergence/divergence assessment, or to attain a high degree of fidelity at the more precise level of estimating years to convergence. As such, it is the only analysis of the four that can be recommended from this study as a simplified analysis without any qualifying statements regarding its situational effectiveness, beyond the warning that its fidelity to the Species Abundance results is not perfect and its performance in qualitative interpretation of NMDS and estimating years to convergence was as inconsistent as the other four analyses. However, this study has highlighted that the effectiveness of simplified analyses can vary even between seemingly similar sites, and so further research should be done to verify this apparent success.

Species Presence/Absence, like the other three analyses, has community compositions under which its performance improves. However, unlike the other

three, Species Presence/Absence analysis has a relatively simple relationship with Species Abundance analysis, as it is a component of Species Abundance. Namely, Species Presence/Absence analysis will replicate the results of Species Abundance analysis more accurately as the differences between site samples becomes increasingly defined by the presence or absence of species at a given site rather than their abundance. That is to say, the higher the species turnover throughout the restoration process at a site, the better Species Presence/Absence analysis will reflect Species Abundance analysis, as the Species Abundance becomes increasingly more like a Species Presence/Absence analysis. Such a scenario is more likely when there are a greater number of specialist species as opposed to generalist species. Both Callide and German Creek mines display Species Presence/Absence-friendly community composition, with less than 5% of species present in more than half the samples collected at both sites, and over 60% of species occur in under 10% of samples. Species Presence/Absence analysis will likely also more closely reflect Species Abundance analysis at sites where the variation in species abundances is lower, such that the relative "weight" of each species in the abundance analysis is more even. The transformation of data to down-weight highly abundant species such as through square-root transformation, as in this study, has the same effect, and, converting abundance data to presence/absence data effectively can be thought of as an extreme down-weighting transformation for abundance data, equalising the abundances of all species so only presence/absence of species affects the analysis (Clarke & Warwick, 2001).

Species Presence/Absence analysis is frequently combined with other simplified analyses in the literature on ants, rather than tested by itself (Pik, Oliver & Beattie, 1999, Andersen *et al*, 2002). However, the use of presence/absence data versus abundance data has been extensively studied in regard to monitoring of freshwater pollution with benthic macroinvertebrates, and found to successfully distinguish between polluted and clean sites and produce multivariate patterns closely resemble those produced by species abundance data (Wright *et al*, 1995, Thorne, Williams & Cao 1999, Marshall, Steward & Harch, 2006). However, it has

been noted that abundance data provides an extra level of discrimination within the "polluted" category between sites of differing water quality, owing to the fact that while the same species were present at both site types, there was variance in the dominance hierarchy of those species (Thorne, Williams & Cao, 1999). There are likely to be similar scenarios in the rehabilitation of mine-sites, where reference and rehabilitated sites are often clearly separated, but patterns within these clusters are also of interest, such as this study. Indeed, the importance of species abundances is demonstrated by the NMDS ordinations of species abundance and species presence/absence at German Creek (fig. 3 (Chapter 2 Results), fig. 16), which show completely different arrangements of rehabilitated sites due to the patterns for species abundance being defined by changes in abundance of two hyperabundant *Iridomyrmex* species. The capacity of Species Presence/Absence analysis to be combined with other simplified analyses offers an intriguing avenue for further simplification of ant community analysis beyond what was explored in this study, and in turn further reduction in time and effort needed per sample. When combined with Large-Bodied Species Abundance analysis and Genus Abundance analysis, as well as family-level analysis, it has been found that little information is lost, although 10% more information was lost when combined with family analysis (Pik, Oliver & Beattie, 1999, Andersen *et al*, 2002, Marshall, Steward & Harch, 2006). However, when combined with Functional Group analysis, the use of log and presence/absence transformations of the data resulted in substantial deviation from the species abundance results (Pik, Oliver & Beattie, 1999). This suggests that, within a degree of separation from species abundance data, simplified analyses may be combined for greater simplification in assessment.

So, under what management contexts is Species Presence/Absence analysis appropriate? Species Presence/Absence analysis, uniquely among the four, appears to be a consistently representative simplified analysis across a variety of contexts, although as with all simplified analyses, there is a loss of discriminatory power, and particularly at fine-scale analysis such as post-rehabilitation mine-site recovery trends, this may cause problems. However, the results of this study

suggest Species Presence/Absence can maintain a high level of fidelity to species abundance results even at this level. As such it appears suited to use in mine-site rehabilitation projects in place of species abundance analysis, although it should be noted that it is not a perfect replacement, as shown by this study and elsewhere. Unfortunately the use of species presence data still necessitates expert identification of the species, but is quicker than species abundance sampling, allowing for more intensive or longer-term sampling, which would hopefully encourage longer studies to capture the full extent of recovery dynamics and allow for quantified assessment of recovery success, rectifying a major issue in mine-site rehabilitation studies to date, as discussed in chapters 2 and 3. The potential combination of simplified analyses for further streamlining and cost-reduction of assessments offers another avenue of research to build upon the results and themes of this study and yield further improvements in the simplification of mine-site assessment.

1.4 – The reliability of simplified analyses

While this study suggests that simplified analyses have potential under the right conditions, it is worth noting that across the literature, other studies have generally found the four simplified analyses trialled here to be even more successful, as indicated in the discussion of each of the four analyses above. Those results are not under dispute, but it should be pointed out that, in general, the level of fidelity sought in other studies tends to be lower than that sought in this study. Most studies of simplified analyses are built around the comparison of two "types" of sites – disturbed and undisturbed – and the ability of simplified analyses to distinguish between them, or evaluate whether simplified analyses are able to track changes in community composition along environmental gradients. In contrast, the focus of this study was on the arrangement of samples within those separated clusters, and as such the criteria for success were naturally more demanding, and even at high levels of correlation in Procrustes analysis, the consistency of rehabilitated-site-specific patterns was quite variable. Unfortunately these small-scale patterns are not easily tested using conventional means of assessment such as Procrustes analysis, the Mantel test

or ANOSIM, particularly when the change over time is taken into account, and there is insufficient data for individual sites to do so in this study in any case. As such, the results of this study cannot be quantitatively compared to results from other studies of simplified analyses, so an explicitly quantitative assessment of how these results fit in with the rest of the literature is not possible.

The implications of this are twofold. First, it means that, for studies such as this, which follow the recovery of rehabilitated mine-sites through time, or indeed any study which tracks changes in community composition over time, the literature paints an overly rosy picture of the utility of simplified analyses, as attested to by the results of this study. This is not to say that simplified analyses are unsuited to long-term studies, only that a higher threshold of fidelity to species abundance results is required for them to be useful. The second implication is that, for most disturbance studies, the results of this study are likely to be over-exacting, or not quite assessing the right elements of simplified analyses for most practitioners. However, I also hope that these results have shown that more and longer multi-year studies can be done at lower cost through the use of simplified analyses, particularly Species Presence/Absence analysis, and invite further study of simplified analyses at this level of precision.

1.5 – Summary Cost-benefit Analysis

Analysis method	Costliness rank	Accuracy at distinguishing reference and rehabilitated sites	Accuracy at identifying trends over time (direction)	Accuracy of identifying trends over time (rate & shape)	Accuracy rank
Species Abundance	**5**	**Very high**	**Very high**	**Very high**	**1**
Genus Abundance	1	Medium	Medium	Low	3
Functional Groups	2	Medium	Medium	Low	3
Large-bodied Abundance	3	Medium	Medium-Low	Low	4
Species Presence/ Absence	4	High	High	Medium	2

Table 22: *A broad summary cost-benefit analysis of the different simplified analyses assessed in this thesis for general utility. Costliness rank is an amalgamation of time and expertise required, and is qualitative assessment based on published commentary (Andersen, 1990, Lattke, 2000, Andersen et al, 2002, Arcoverde et al, 2017). Accuracy measures are judged based on combined performance at Callide & German Creek mines. Accuracy rank is a combination of the three accuracy measures. Accuracy measures use Species Abundance analysis results as a baseline, with assumption Species Abundance analysis is "correct". Which analysis should actually be used will depend on the cost restrictions and accuracy priorities of the practitioner and on the site ecology.*

2 – Conclusions

The use of simplified analyses, and which to use, is a topic that continues to attract a lot of interest and be heavily debated. These results contribute to that mixed picture, indicating that the successful use of simplified analysis depends on factors including, but not limited to, the ecology of the study site, the strength of discrimination sought, and the management context in which they are applied. Thus, while the rewards may be rich when used successfully in terms of time and funding saved, using them without careful consideration can be very risky. Of the four simplified analyses tested in this study, only Species Presence/Absence analysis emerges as a strong candidate for a consistently reliable, readily usable simplified analysis. The conclusions and conjectures of this study are also somewhat constrained by the limited sample size. Nonetheless, the take-home message of this work is that all four of these simplified analyses can, under the right conditions, be a very successful alternative to a full species-abundance analysis, and all offer substantial reductions in time, effort and cost required for mine-site assessment, making these conditions well worth further study. This study has also been a trial of the statistical techniques through which we can explore these questions, and although they remain subject to a variety of practical constraints, ARC-Dissimilarity remains a conceptually strong approach to tracking disturbed or rehabilitated-site community change under the convergence model, which shows good complementarity with standard multivariate ordination techniques and a potentially powerful and intriguing synergy with the use of simplified analyses, showing greater consistency in congruence to species abundance results than NMDS results for the same analysis. Success in these endeavours could open up the field to more in-depth study and assessment of mine-site rehabilitation, and in doing so improve our understanding and application of rehabilitation and restoration to mine sites and other degraded ecosystems. As such, this study should be seen as a jumping-off point for further study of the use of simplified analyses for time-series analysis and site-specific trends, and of the underlying mechanisms in ant community composition and statistical methods that determine when and where simplified analyses can really shine.

Funding for ecological fieldwork is under pressure, and it is essential for surveys to maximise the information extracted for each unit of input, in terms of time, expertise and ultimately, money. This funding pressure coincides with a period of unprecedented pressure on ecosystems and biodiversity as a result of human activity, to the point that the current time has been declared "The Anthropocene" due to the extent of human influence on the globe's ecosystems (Crutzen & Stoermer, 2000), and the current rate of biodiversity loss has been found to be comparable to those of the Big Five mass extinction events (Barnosky *et al*, 2011). The need for robust biological monitoring to ascertain the extent of the damage and inform conservation and restoration interventions, both at the broad, global scale of threats like climate change and acute disturbances such as mining, is greater than ever, and yet we are suffering not only through a shortage of funding but a shortage of data – 10% of species assessed by the IUCN are classed as Data Deficient, when there is "inadequate information to make a direct, or indirect, assessment of its risk of extinction based on its distribution and /or population status" (IUCN, 2001, Kindsvater *et al*, 2018). As such, we must not only find ways to maintain our current level of monitoring, but expand it significantly, in spite of funding shortfalls, if we hope to meet the ecological challenges of the 21st century. There is hope – there is increasing public and governmental support and pressure for sustainable, eco-friendly development – but without accurate monitoring data, it is difficult, if not impossible, to gain support for and acceptance of regulations, set reasonable targets, or to evaluate and enforce them (Majer & Nichols, 1998, Hamburg *et al*, 2004, Kindsvater *et al*, 2018). But it is not generally feasible to census an entire ecosystem, especially repeatedly to monitor the effects of environmental disturbance and other threats. Bioindicators then are fundamental to making environmental monitoring feasible, allowing us to extract the information we require from a small, relatively easily surveyed component of complex systems. But as we have seen, even when we have access to representative, relatively easily sampled taxa to use as bioindicators, this is often still insufficient to achieve fully comprehensive, long-term monitoring, which remains rare, even for

relatively small-scale disturbances, let alone for challenges on the global scale (Andersen & Majer, 2004). This not only constrains the vital task of monitoring itself, but also hampers our ability to study the ecology of disturbance and recovery, and of biomonitoring, preventing us from making the advances we need to address our deficits in monitoring, conservation and rehabilitation of degraded ecosystems. So, to address this, we must follow the conceptual basis of bioindicators to its logical conclusion and find means to further streamline the process of ecological monitoring, making it quicker, making it cheaper, making it easier to monitor our chosen indicators. One of the keys to this challenge is finding viable simplified analyses, finding the minimum amount of information we need on a given system in order to accurately monitor the environmental variables we are interested in, be it pollution levels, rates of biodiversity change or ecosystem function provision. This minimum will vary depending on the focal variables, ecosystems and indicators, as well as monitoring objectives, and so, while at some levels we may expect to find generally reliable simplified analyses, and such analyses should be highly prized, the search for simplified analyses should be largely considered a context-driven enterprise. With an acknowledgement and understanding of how the correlation of proposed simplified analyses such as higher taxonomic resolutions or ecologically-defined groupings with species abundance varies under different conditions (Andersen, 1995b), we can adopt a more nuanced approach to the identification and deployment of simplified analyses. We can then move past circular debates over the effectiveness of particular approaches (Souza *et al*, 2018) and instead focusing our efforts on establishing why such approaches work well when they do, allowing us to optimise our monitoring efficiency under a given set of circumstances with conditionally-effective analyses rather than being forced to rely on globally consistently reliable but contextually suboptimal analytical approaches. We must do all we can to maximise the efficiency of ecological monitoring, as under the current paradigm we lack the resources to adequately achieve our goals. But with the use of simplified analyses, we have the potential to greatly improve the efficiency with which we can leverage those resources,

and in doing so dramatically amplify our capacity to monitor and maintain the global environment.

Appendix 1 – Functional Group Relative Abundance Profiles

Section 1 – Callide Sites

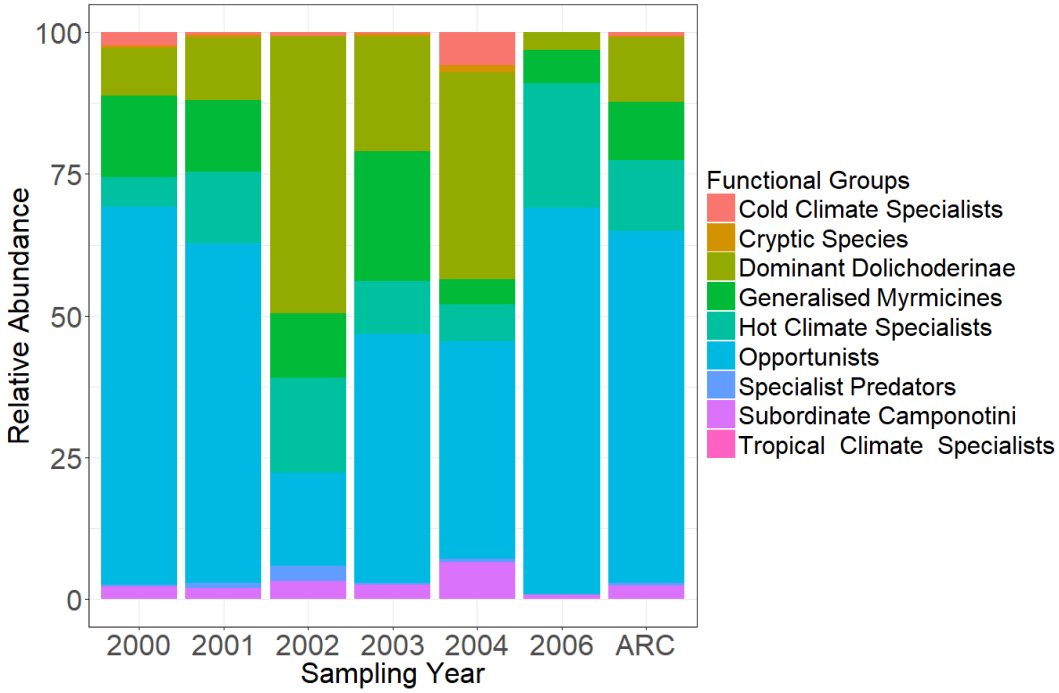


Figure 21: *Functional Group relative abundance profile for Callide Reference Site 6. Profile shows relative abundances of each functional group in each community sample from the site and in the Average Reference Community (ARC).*

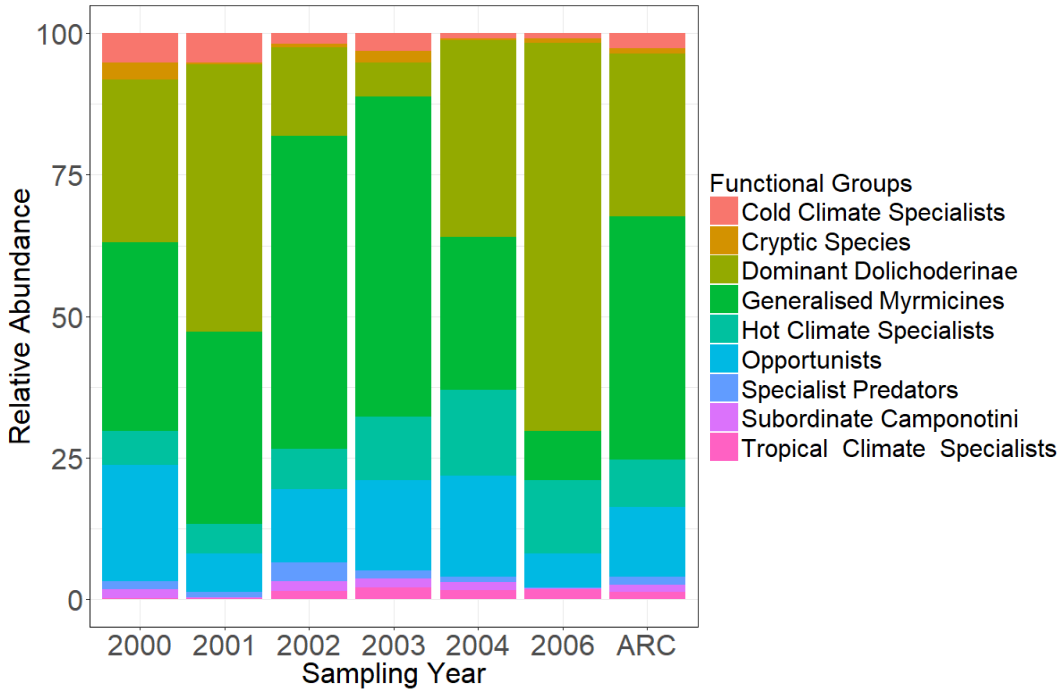


Figure 22: *Functional Group relative abundance profile for Callide Reference Site 8. Profile shows relative abundances of each functional group in each community sample from the site and in the Average Reference Community (ARC).*

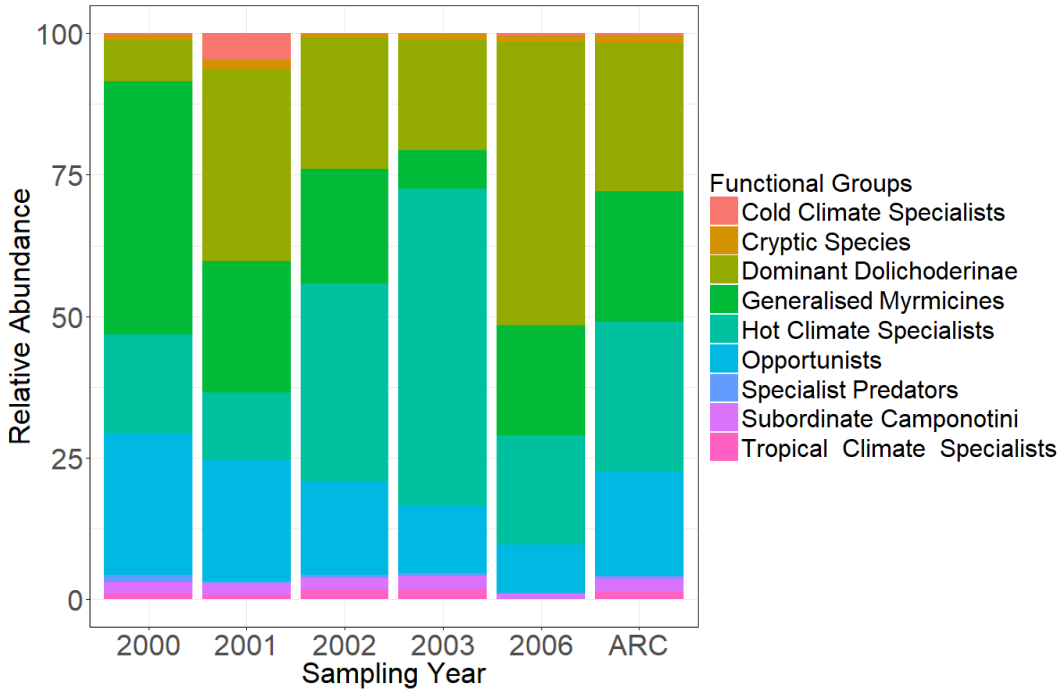


Figure 23: *Functional Group relative abundance profile for Callide Reference Site 9. Profile shows relative abundances of each functional group in each community sample from the site and in the Average Reference Community (ARC).*

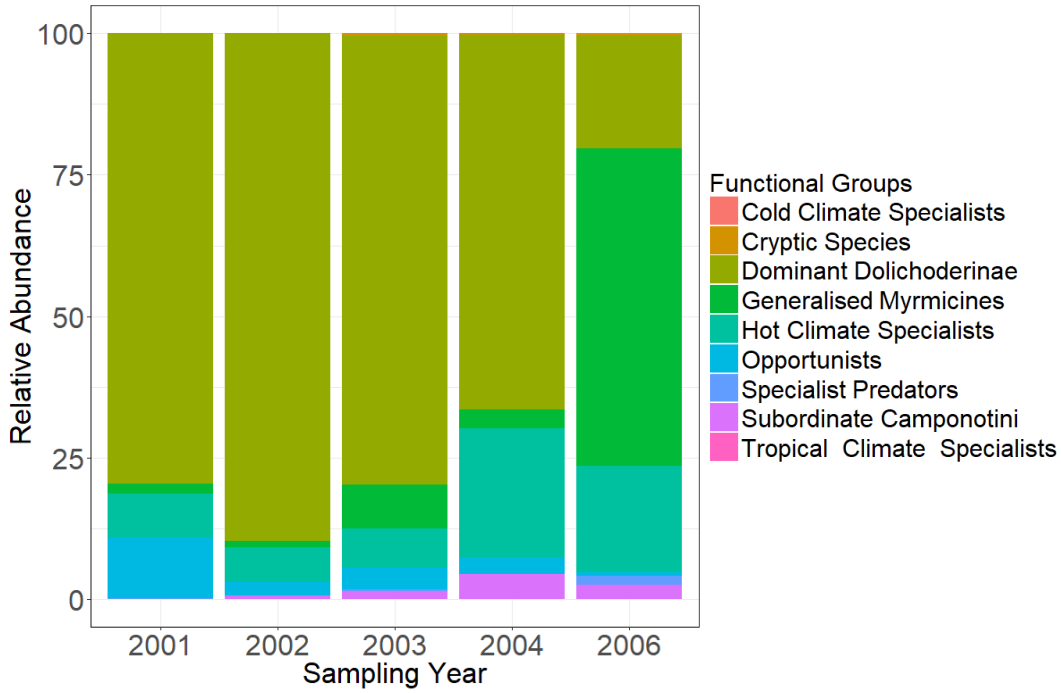


Figure 24: *Functional Group relative abundance profile for Callide Rehabilitated Site BH99. Profile shows relative abundances of each functional group in each community sample from the site.*

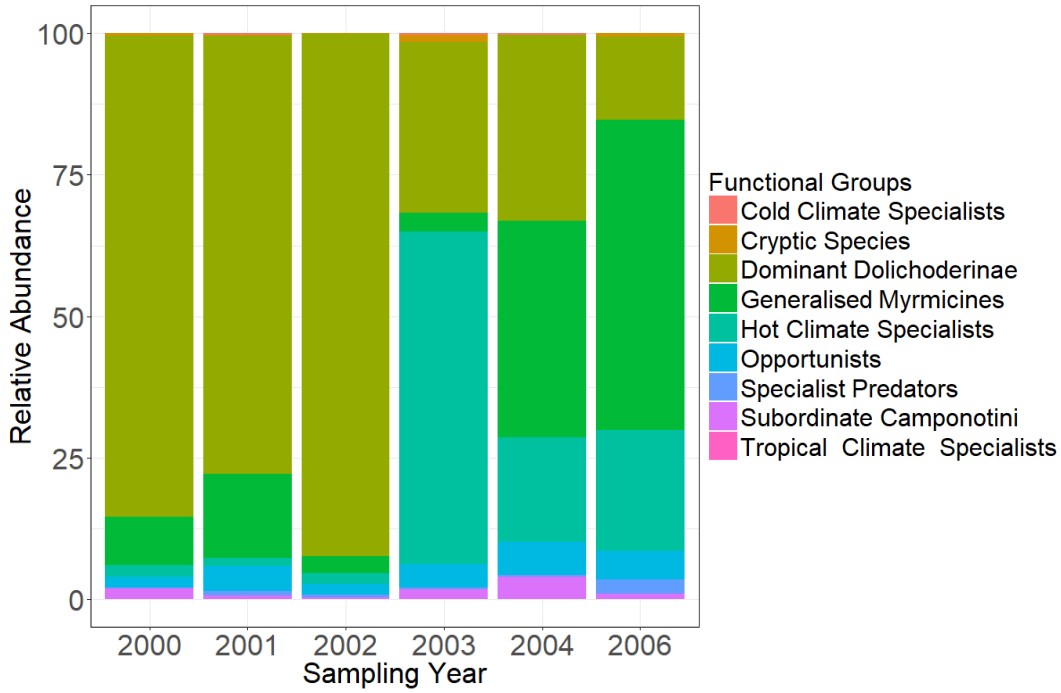


Figure 25: *Functional Group relative abundance profile for Callide Rehabilitated Site BSH94. Profile shows relative abundances of each functional group in each community sample from the site.*

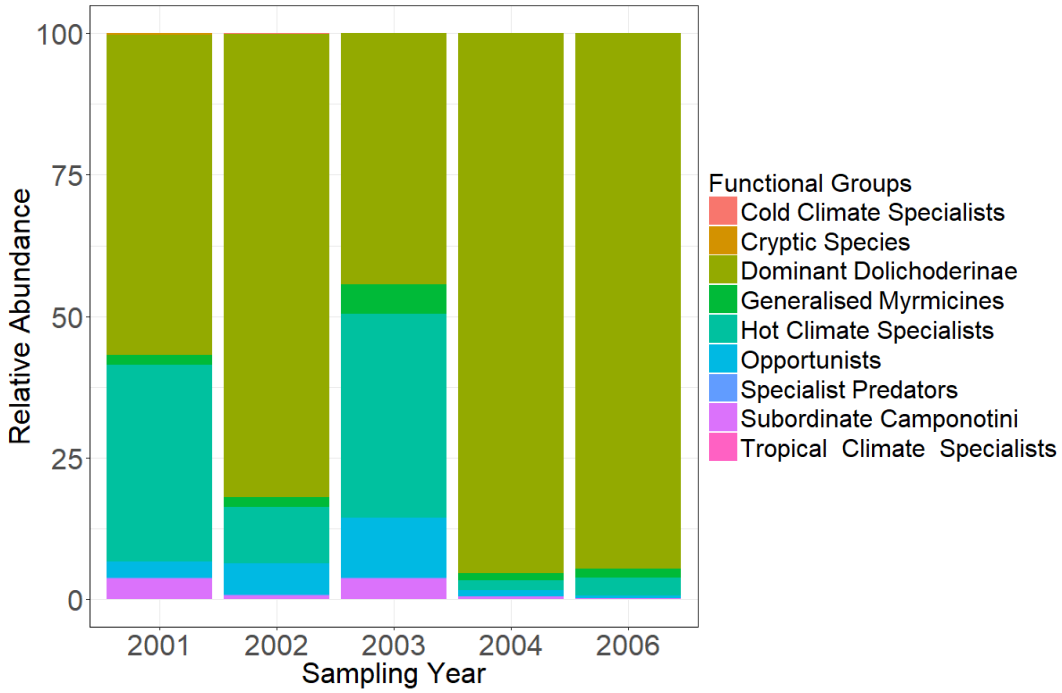


Figure 26: *Functional Group relative abundance profile for Callide Rehabilitated Site DCB94. Profile shows relative abundances of each functional group in each community sample from the site.*

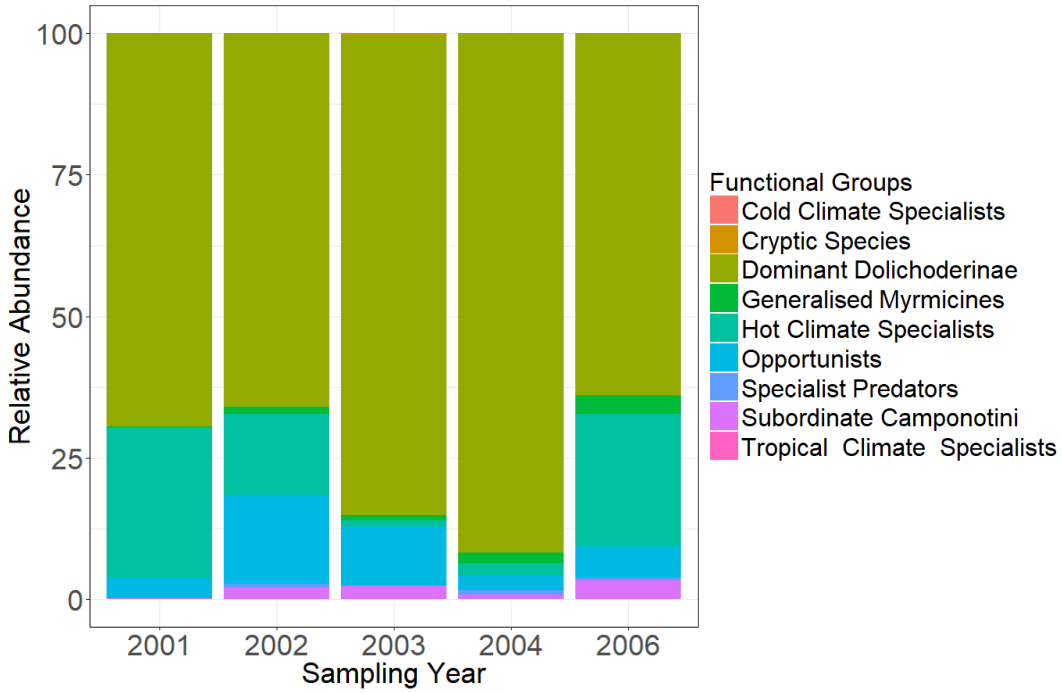


Figure 27: *Functional Group relative abundance profile for Callide Rehabilitated Site DCB98. Profile shows relative abundances of each functional group in each community sample from the site.*

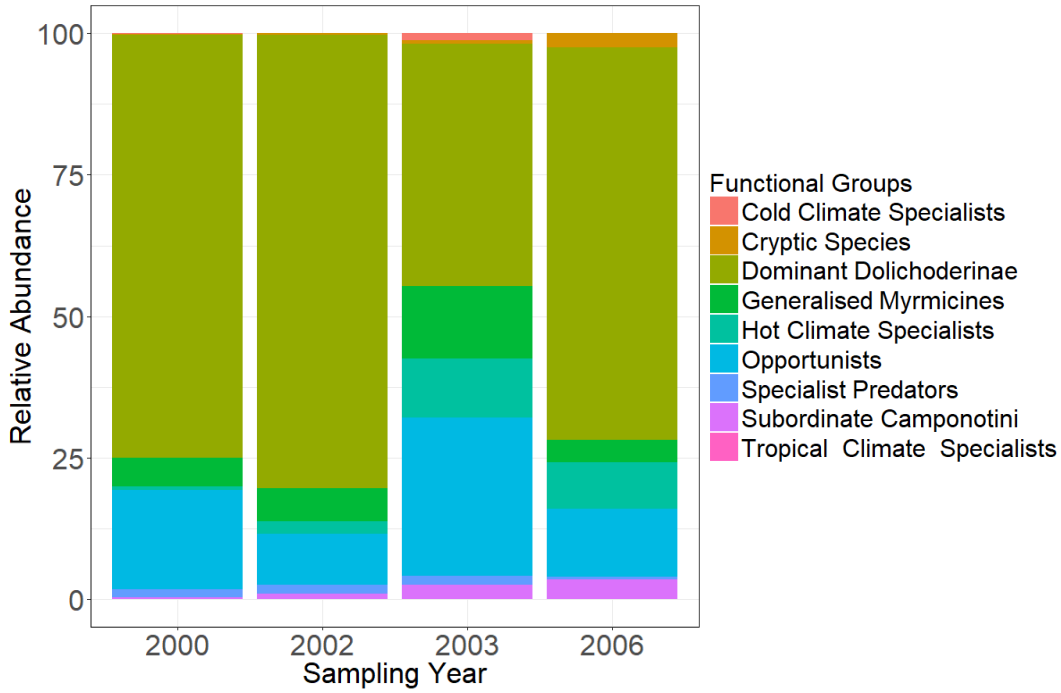


Figure 28: *Functional Group relative abundance profile for Callide Rehabilitated Site DSC81. Profile shows relative abundances of each functional group in each community sample from the site.*

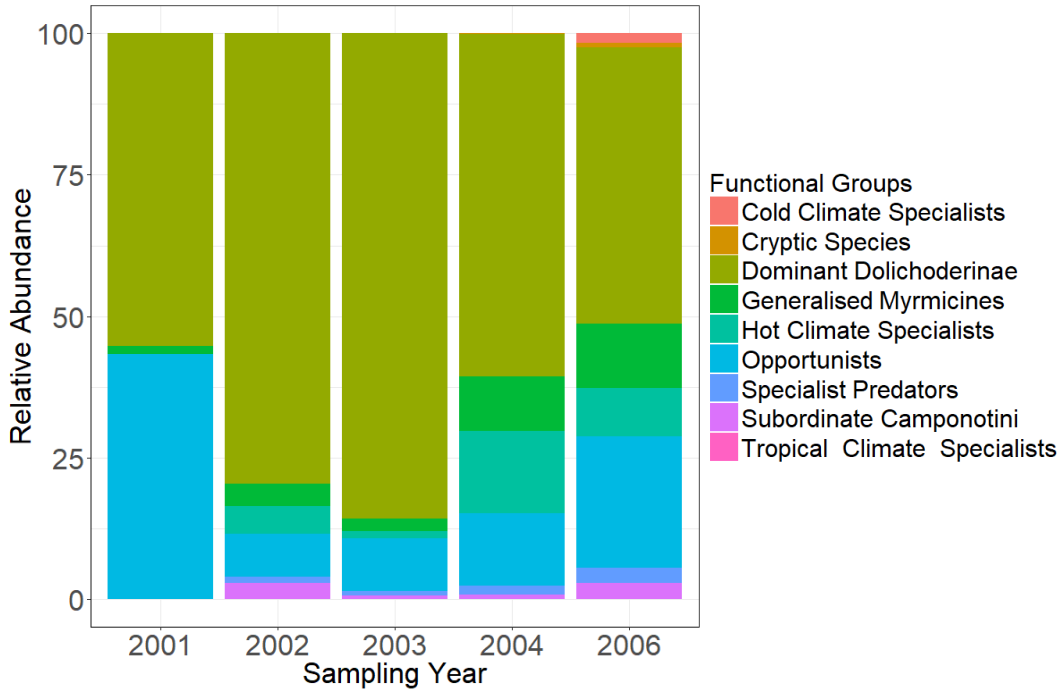


Figure 29: *Functional Group relative abundance profile for Callide Rehabilitated Site TGB98. Profile shows relative abundances of each functional group in each community sample from the site.*

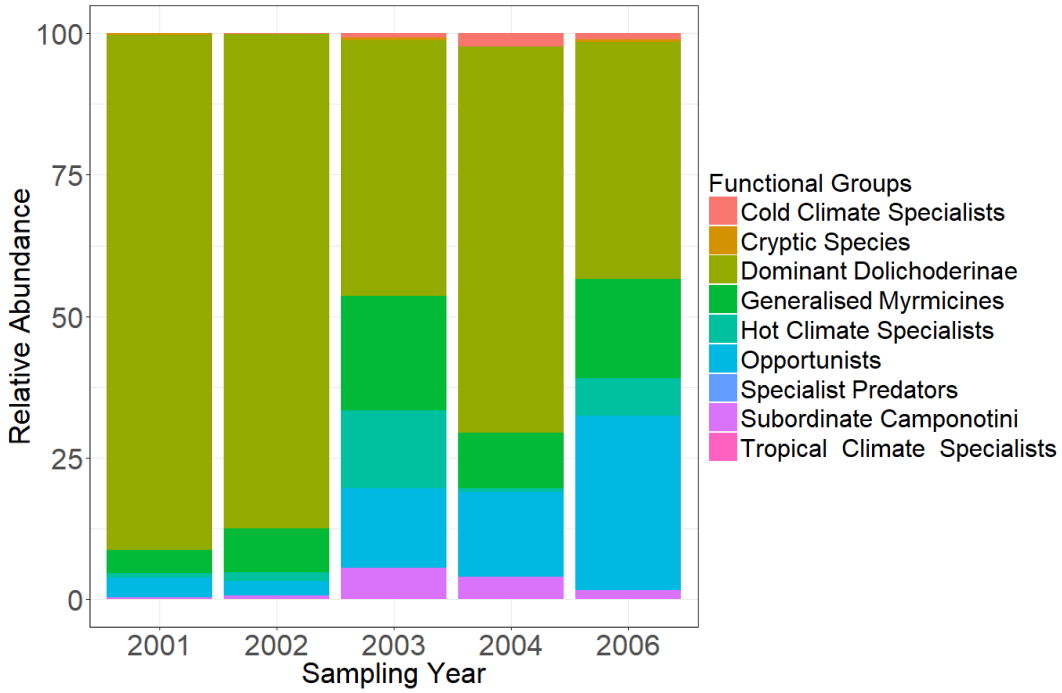


Figure 30: *Functional Group relative abundance profile for Callide Rehabilitated Site TGC92. Profile shows relative abundances of each functional group in each community sample from the site.*

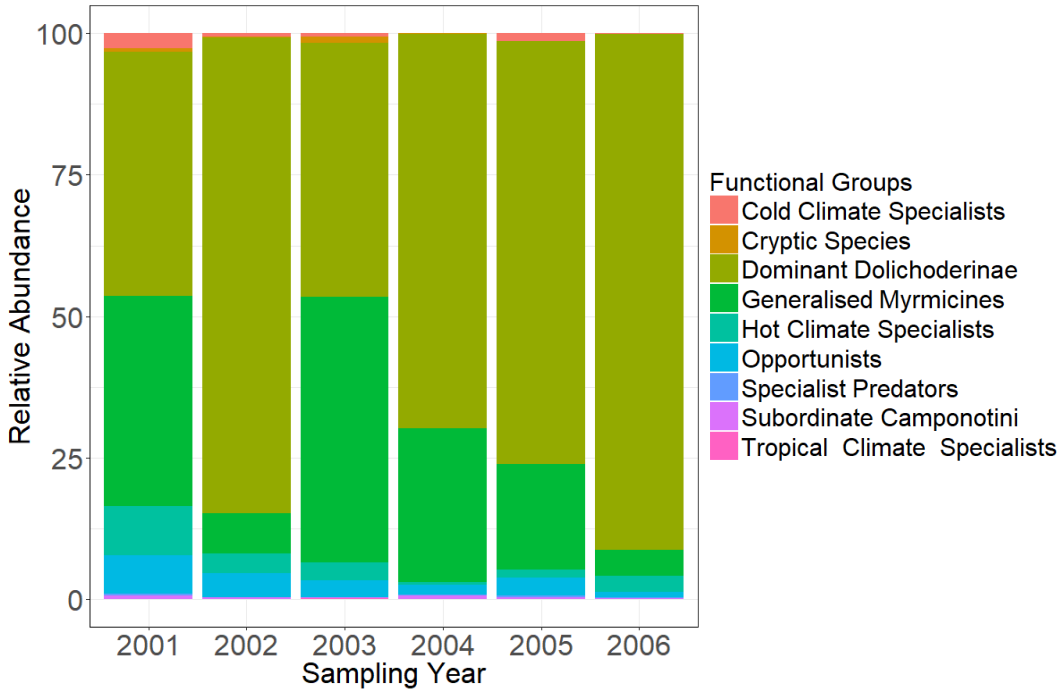


Figure 31: *Functional Group relative abundance profile for Callide Rehabilitated Site TH91. Profile shows relative abundances of each functional group in each community sample from the site.*

Section 2 – German Creek Sites

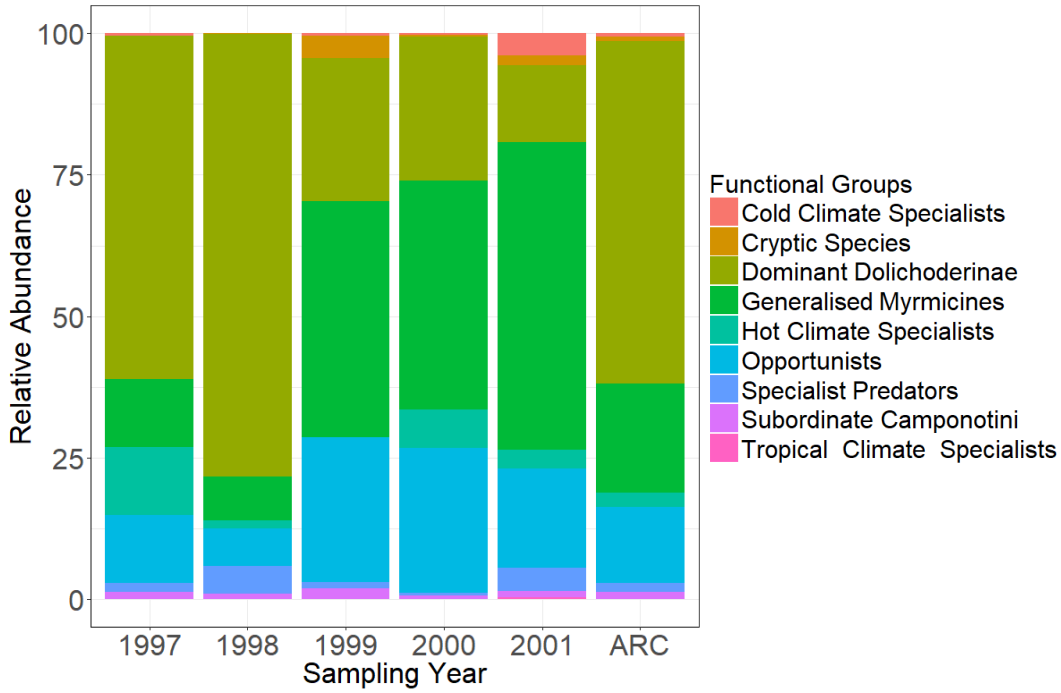


Figure 32: *Functional Group relative abundance profile for German Creek Reference Site 1. Profile shows relative abundances of each functional group in each community sample from the site and in the Average Reference Community (ARC).*

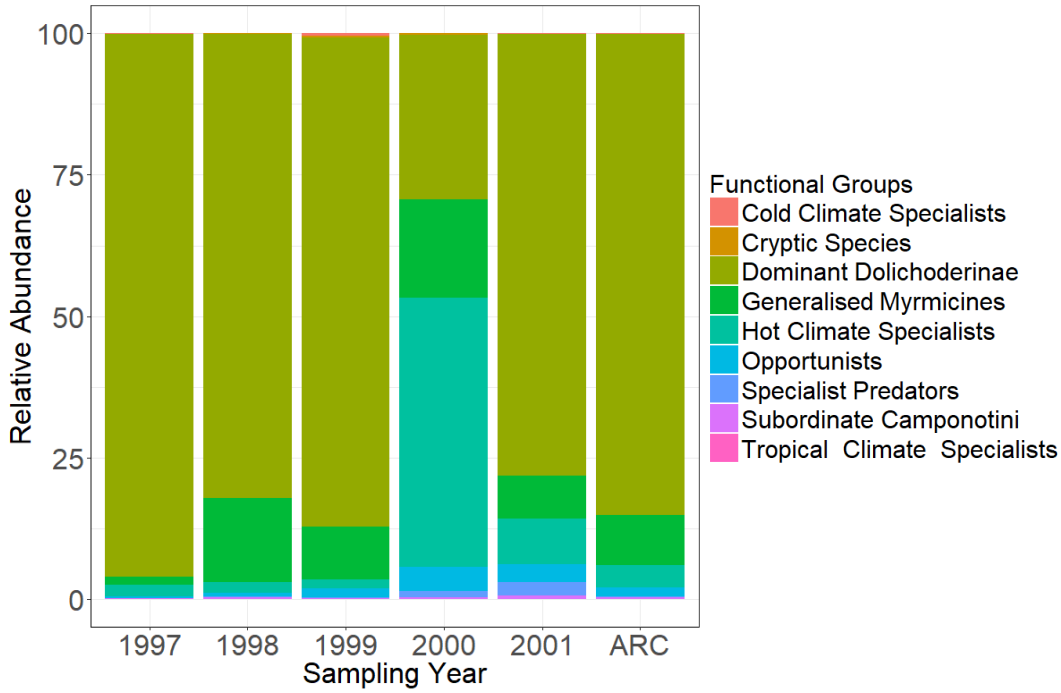


Figure 33: *Functional Group relative abundance profile for German Creek Reference Site 3. Profile shows relative abundances of each functional group in each community sample from the site and in the Average Reference Community (ARC).*

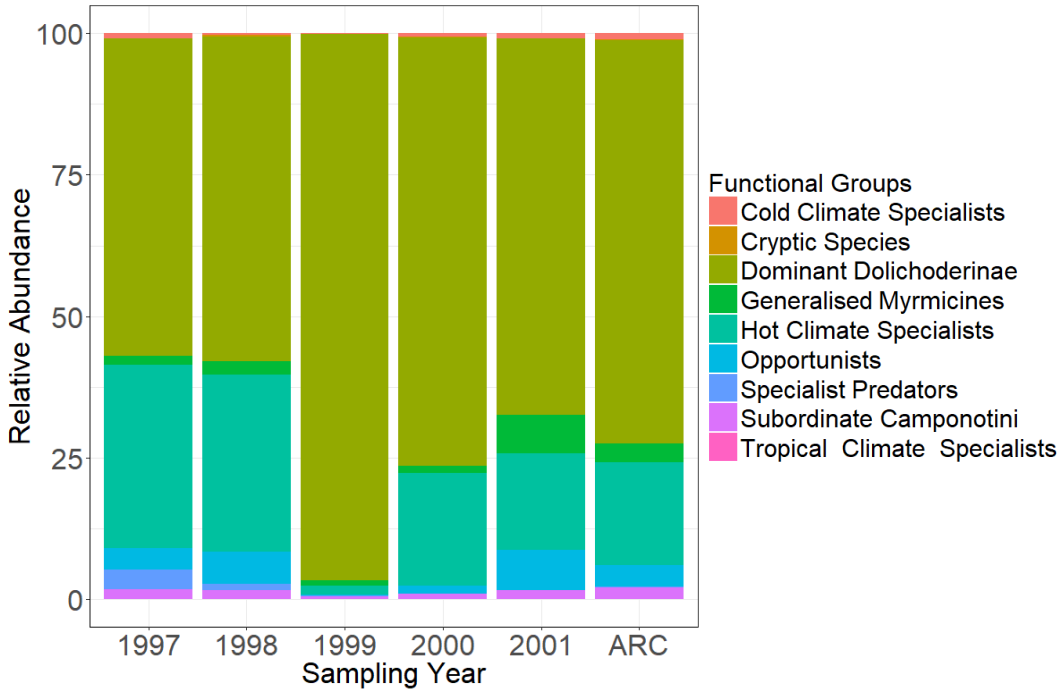


Figure 34: *Functional Group relative abundance profile for German Creek Reference Site 7. Profile shows relative abundances of each functional group in each community sample from the site and in the Average Reference Community (ARC).*

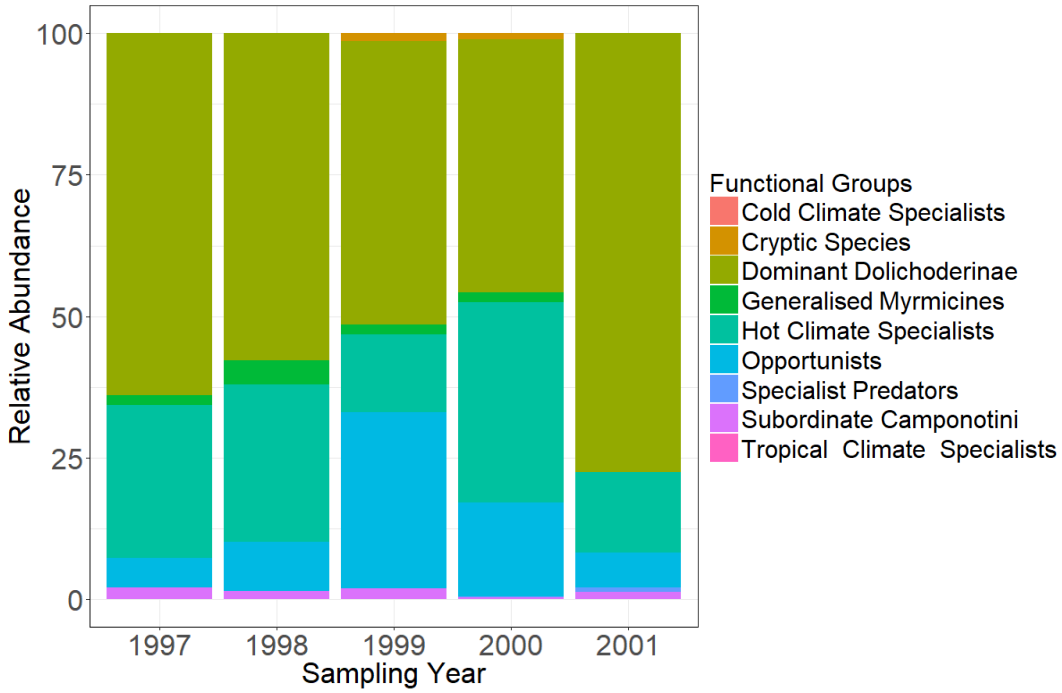


Figure 35: *Functional Group relative abundance profile for German Creek Rehabilitated Site 8. Profile shows relative abundances of each functional group in each community sample from the site.*

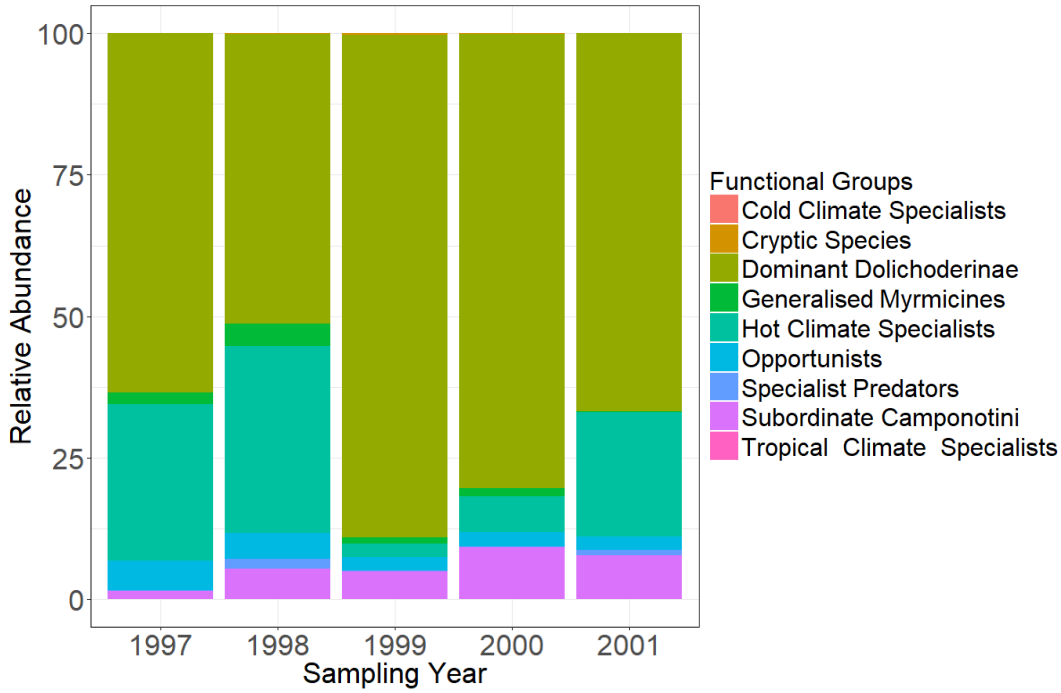


Figure 36: *Functional Group relative abundance profile for German Creek Rehabilitated Site 9. Profile shows relative abundances of each functional group in each community sample from the site.*

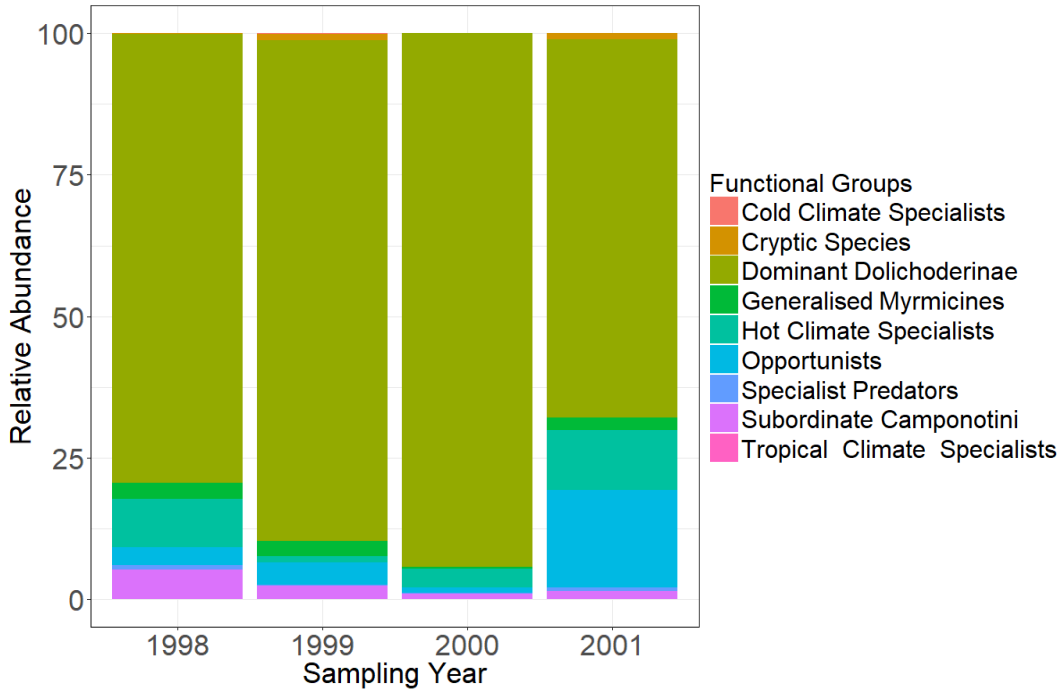


Figure 37: *Functional Group relative abundance profile for German Creek Rehabilitated Site 10. Profile shows relative abundances of each functional group in each community sample from the site.*

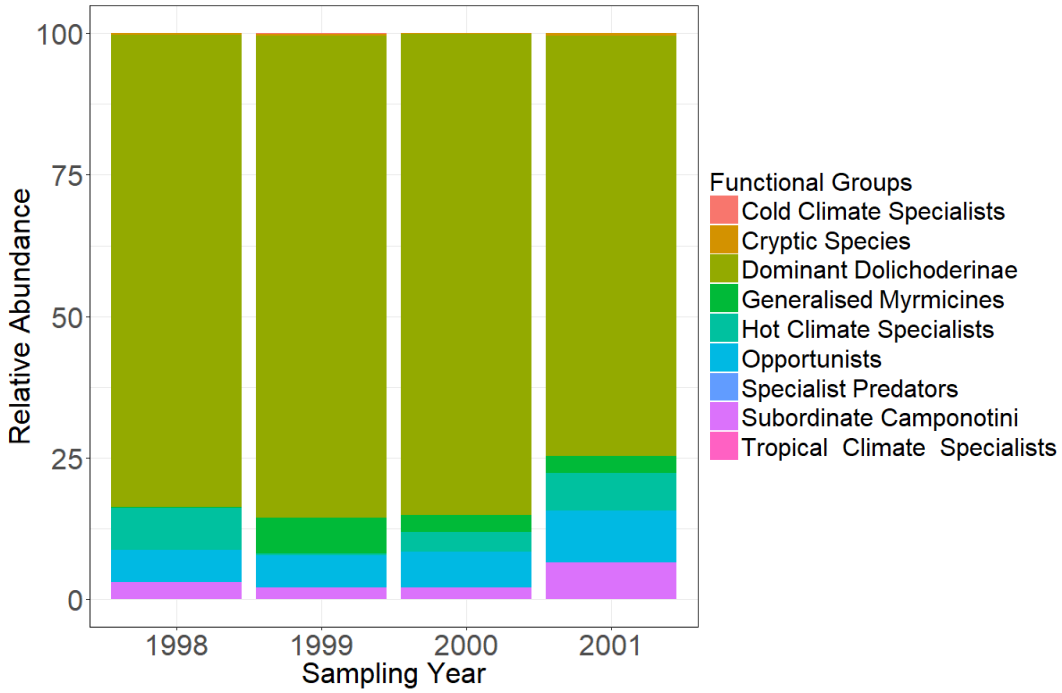


Figure 38: *Functional Group relative abundance profile for German Creek Rehabilitated Site 11. Profile shows relative abundances of each functional group in each community sample from the site.*

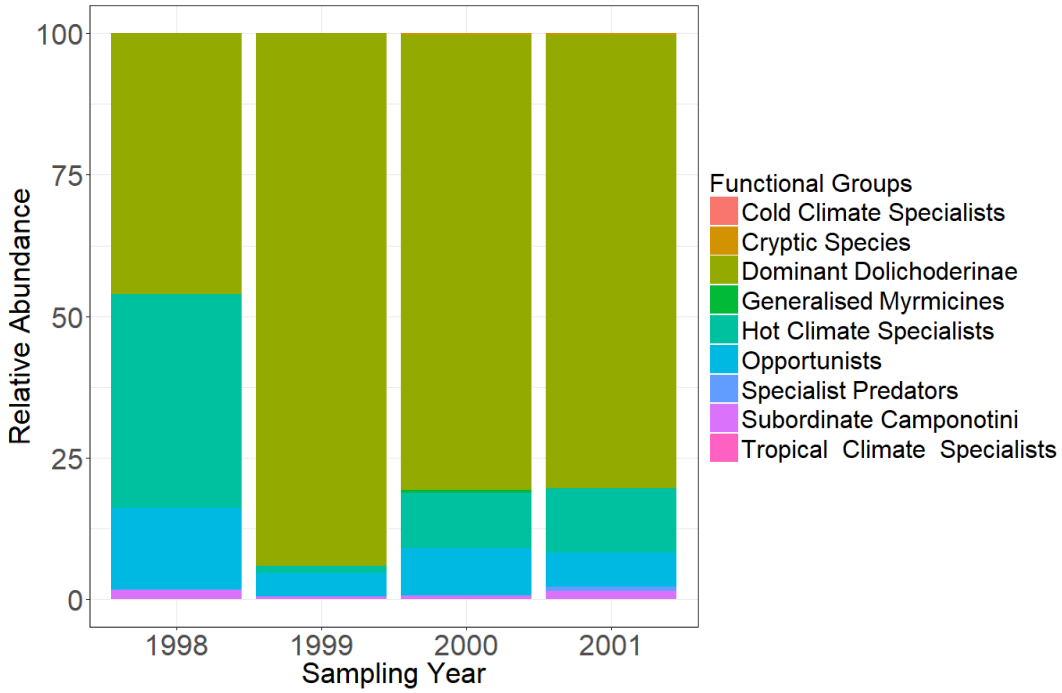


Figure 39: *Functional Group relative abundance profile for German Creek Rehabilitated Site 12. Profile shows relative abundances of each functional group in each community sample from the site.*

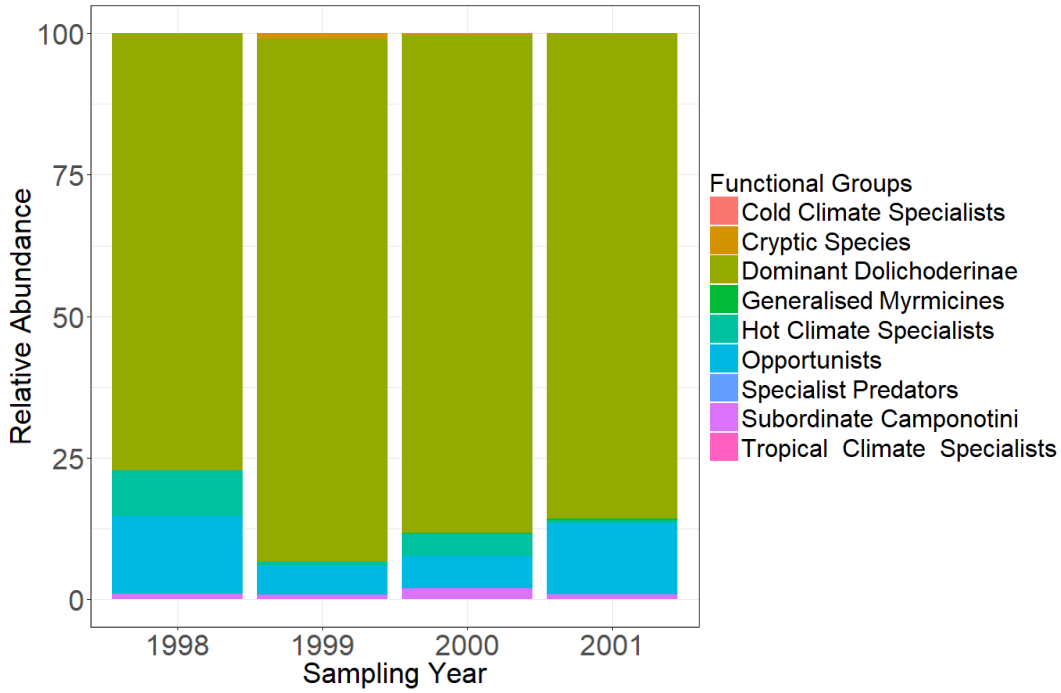


Figure 40: *Functional Group relative abundance profile for German Creek Rehabilitated Site 13. Profile shows relative abundances of each functional group in each community sample from the site.*

Appendix 2 – Reciprocal Transformation of Estimated Years to Convergence

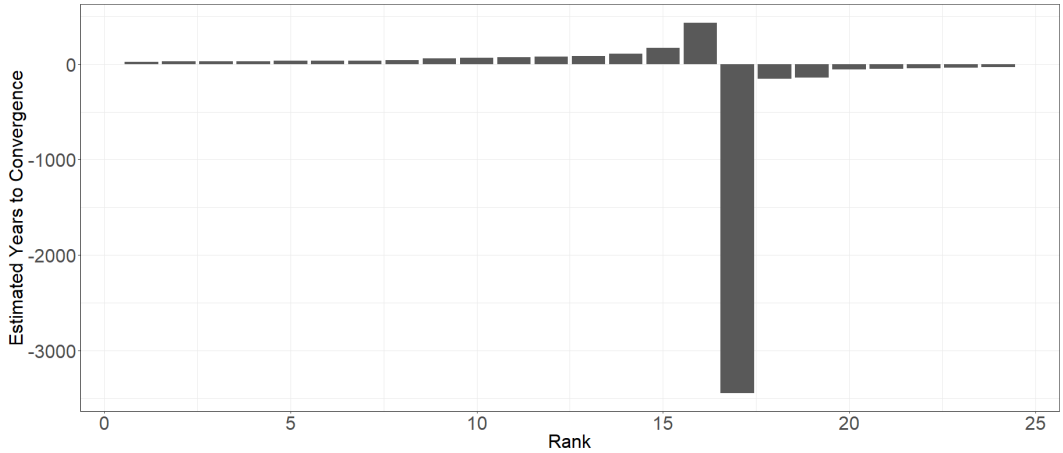


Figure 41: *Estimated Years to Convergence (EYCA) of each rehabilitated site to each of the three reference sites at Callide Mine, based on Species Abundance data. Estimates have been ranked from most to least convergent, or least to most divergent. Positive values are convergent, negative values are divergent.*

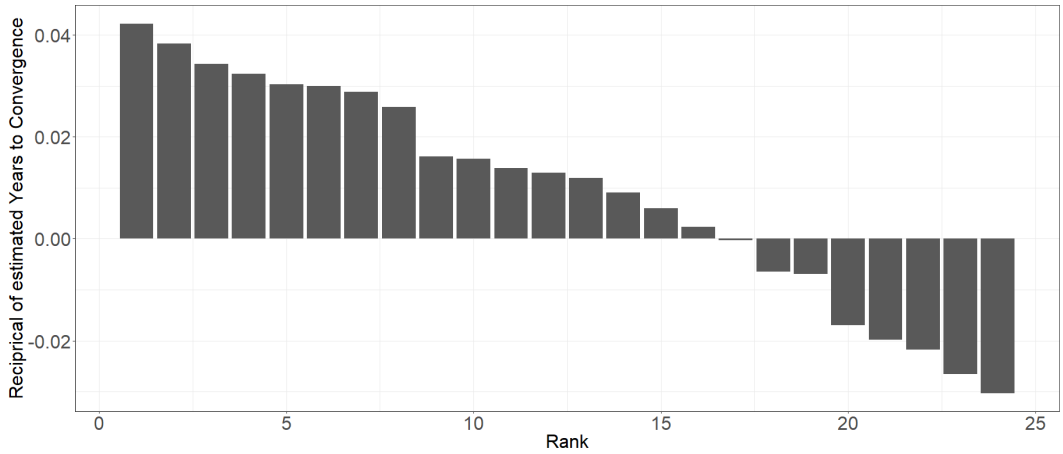


Figure 42: *Reciprocals of Estimated Years to Convergence (EYCA) of each rehabilitated site to each of the three reference sites at Callide Mine, based on Species Abundance data. Estimates have been ranked from most to least convergent, or least to most divergent. Positive values are convergent, negative values are divergent.*

References

- Agosti, D. & Alonso, L.E. (2000). The ALL protocol: a standard protocol for the collection of ground-dwelling ants. In Agosti, D., Majer, J.D., Alonso, L.E. & Schultz, T.R. (Eds.) *ANTS: Standard methods for measuring and monitoring biodiversity*, Smithsonian Institute Press, Washington, 204-206.
- Andersen, A.N. (1983). Species diversity and temporal distribution of ants in the semi-arid mallee region of northwestern Victoria. Australian Journal of Ecology, 8, 127-137.
- Andersen, A.N. (1990). The use of ant communities to evaluate change in Australian terrestrial ecosystems: a review and a recipe. *Proceedings of the Ecological Society of Australia*, 16, 347-357.
- Andersen, A.N. (1993). Ants as indicators of restoration success at a uranium mine in tropical Australia. *Restoration Ecology*, 1(3), 156-167.
- Andersen, A.N. (1995a). A classification of Australian ant communities, based on functional groups which parallel plant life-forms in relation to stress and disturbance. *Journal of Biogeography*, 22(1), 15-29.
- Andersen, A.N. (1995b). Measuring more of biodiversity: genus richness as a surrogate for species richness in Australian ant faunas. *Biological Conservation*, 73, 29-43.
- Andersen, A.N. (1997). Functional groups and patterns of organisation in North American ant communities: a comparison with Australia. *Journal of Biogeography*, 24, 433-460.
- Andersen, A.N. (2000). *The Ants of Northern Australia: A Guide to the Monsoonal Fauna*. CSIRO Publishing, Collingwood, Australia.
- Andersen, A.N. & Hoffmann, B.D. (1998a). *Ant Monitoring at German Creek Mine, Queensland*. A consultancy report prepared for Capricorn Management Pty Ltd. CSIRO Wildlife and Ecology, Darwin.
- Andersen, A.N. & Hoffmann, B.D. (1998b). *Ant Monitoring at German Creek Mine, QLD. 1998 Sampling*. A consultancy report prepared for Capricorn Coal Management Pty Ltd. CSIRO Wildlife and Ecology, Darwin.

- Andersen, A.N. & Hoffmann, B.D. (1999). *Ant Monitoring at German Creek Mine, QLD. 1999 Sampling*. A consultancy report prepared for Capricorn Coal Management Pty Ltd. CSIRO Wildlife and Ecology, Darwin.
- Andersen, A.N. & Hoffmann, B.D. (2000). *Ant Monitoring at German Creek Mine, QLD. 2000 Sampling*. A consultancy report prepared for Capricorn Coal Management Pty Ltd. CSIRO Wildlife and Ecology, Darwin.
- Andersen, A.N. & Hoffmann, B.D. (2001a). *Ant Monitoring at German Creek Mine, QLD. 2001 Sampling*. A consultancy report prepared for Capricorn Coal Management Pty Ltd. CSIRO Sustainable Ecosystems, Darwin.
- Andersen, A.N. & Hoffmann, B.D. (2001b). *Ant Monitoring at Callide Mine, Central Queensland. 2001 Sampling*. A consultancy report prepared for Callide Mine. CSIRO Sustainable Ecosystems, Darwin.
- Andersen, A.N. & Hoffmann, B.D. (2002). *Ant Monitoring at Callide Mine, Central Queensland. 2002 Sampling*. A consultancy report prepared for Callide Mine. CSIRO Sustainable Ecosystems, Darwin.
- Andersen, A.N. & Hoffmann, B.D. (2003a). Responses of ants to disturbance in Australia, with particular reference to functional groups. *Austral Ecology*, 28, 444-464.
- Andersen, A.N. & Hoffmann, B.D. (2003b). *Ant Colonisation of Rehabilitated Mines at German Creek Mine: A Review*. A consultancy report prepared for Anglo Coal. CSIRO Wildlife and Ecology, Darwin.
- Andersen, A.N. & Hoffmann, B.D. (2005). *Ant Monitoring at Callide Mine, Central Queensland. 2005 Sampling*. A consultancy report prepared for Callide Mine. CSIRO Sustainable Ecosystems, Darwin.
- Andersen, A.N. & Hoffmann, B.D. (2006). *Ant Monitoring at Callide Mine, Central Queensland. 2006 Sampling*. A consultancy report prepared for Callide Mine. CSIRO Sustainable Ecosystems, Darwin.
- Andersen, A.N., Hoffmann, B.D., Müller, W.J. & Griffiths, A.D. (2002). Using ants as bioindicators in land management: simplifying assessment of ant community species. *Journal of Applied Ecology*, 39, 8-17.

- Andersen, A.N., Hoffmann, B.D., Somes, J. (2003). Ants as indicators of minesite restoration: community recovery at one of eight rehabilitation sites in central Queensland. *Ecological Management & Restoration*, 4(Supplement), S12-S19.
- Andersen, A.N. & Majer, J. D. (2004). Ants show the way down under: invertebrates as bioindicators in land management. *Frontiers in Ecology and Environment*, 2(6), 291-298.
- Andersen, A.N., Fischer, A., Hoffmann, B.D., Read, J.L. (2004). Use of terrestrial invertebrates for biodiversity monitoring in Australian rangelands, with particular reference to ants. *Austral Ecology*, 29, 87-92.
- Andersen, A.N., Ribbons, R.R., Petit, M. & Parr, C.L. (2014). Burning for biodiversity: highly resilient ant communities respond only to strongly contrasting fire regimes in Australia's seasonal tropics. *Journal of Applied Ecology*, 51, 1406-1413.
- Anglo American (2018a) German Creek profile on MiningLink.com.au (accessed 11/01/2018 from http://mininglink.com.au/site/german-creek).
- Anglo American (2018b) Callide profile on MiningLink.com.au (accessed 11/01/2018 from http://mininglink.com.au/site/callide).
- Arcoverde, G., Andersen, A.N., Inara, L. & Setterfield, S. (2017). 'Habitat-contingent responses to disturbance: impacts of cattle grazing on ant communities vary with habitat complexity'. [Unpublished manuscript].
- Bakeman, R., MacArthur, D., Quera, V., Robinson, B. (1997). Detecting sequential patterns and determining their reliability with fallible observers. Psychological Methods, 2(4), 357-370.
- Banning, N.C., Grant, C.D., Jones, D.L. & Murphy, D.V. (2008). Recovery of soil organic matter, organic matter turnover and nitrogen cycling in a post-mining forest rehabilitation chronosequence. *Soil Biology & Biochemistry*, 40, 2021-2031.
- Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quential, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C.,

Mersey, B. & Ferrer, E.A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature*, 471, 51-57.

- Bell, L.C. (2001). Establishing of native ecosystems are mining – Australian experience across diverse biogeographic zones. *Ecological Engineering*, 17, 179-186.
- Bestelmeyer, B.T., Agosti, D., Alonso, L.E., Brandão, Brown Jr., W.L., Delabie, J.H.C. & Silvestre, R. (2000). Field techniques for the study of ground-dwelling ants: an overview, description and evaluation. In Agosti, D., Majer, J.D., Alonso, L.E. & Schultz, T.R. (Eds.) *ANTS: Standard methods for measuring and monitoring biodiversity*, Smithsonian Institute Press, Washington, 122-144.
- Bisevac, L. & Majer, J.D. (1999) Comparative study of ant communities of rehabilitated mineral sand mines and heathland, Western Australia. *Restoration Ecology*, 7(2), 117-126.
- Bisevac, L. & Majer, J.D. (2002). Cost effectiveness and data-yield of biodiversity surveys. *Journal of the Royal Society of Western Australia*, 85, 129-132.
- Briese, D.T. & Macauley, B.J. (1980). Temporal structure of an ant community in semi-arid Australia. Australian Journal of Ecology, 5, 121-134.
- Bureau of Meteorology (2014) Bureau of Meteorology weather stations - All Australia. Bioregional Assessment Source Dataset. Viewed 17 & 18 October 2018
- Burger, J. (2006). Bioindicators: a review of their use in the environmental literature 1970-2005. *Environmental Bioindicators*, 1, 136-144.
- Buttigieg, P.L. & Ramette, A. (2014). A guide to statistical analysis in microbial ecology: a community-focused, living review of multivariate data analyses. *FEMS Microbiology Ecology*, 90, 543-550.
- Cairns, J Jr & Pratt, J.R. (1993). A history of biological monitoring using benthic macroinvertebrates. In Rosenberg, D.M., Resh & V.H. (Eds.).

Freshwater biomonitoring and Benthic Macroinvertebrates. Chapman & Hall, New York, 10-27.

- Clarke, K.R. & Warwick, R.M. (2001). *Changes in Marine Communities: An approach to statistical analysis and interpretation. 2nd Edition.* PRIMER-E Ltd, Plymouth Marine Laboratory.
- Clarke, K.R., Somerfield, P.J. & Chapman, M.G. (2006). On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages. *Journal of Experimental Marine Biology and Ecology*, 330, 55-80.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37-46.
- Cole, L.E.S., Bhagwat, S.A. & Willis, K.J. (2014). Recovery and resilience of tropical forests after disturbance. *Nature Communications*, DOI: 10.1038/ncomms4906
- Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P. & Linder, H.P. (2009). Phylogenetic biome conservatism on a global scale. *Nature*, 458, 754-756.
- Crutzen, P.J. & Stoermer, E.F. (2000). The "Anthropocene". *Global Change Newsletter*, 41, 17-18.
- Cuccovia, A. & Kinnear, A. (1999). Acarine (mite) communities colonizing rehabilitated bauxite mine pits in the jarrah forest of Western Australia. In Ponder, W. & Lunney, D. (Eds.) *The Other 99%: The conservation and biodiversity of invertebrates*, Royal Zoological Society of New South Wales, Sydney, 54-59.
- Dunn, R.R., Agosti, D., Andersen, A.N., Arnan, X., Bruhl C.A, Cerda, X., Ellison, A.M et al (2009). Climatic drivers of hemispheric asymmetry in global patterns of ant species richness. Ecology Letters, 12(4), 324-333.
- Fox, B.J. (1982). The influence of disturbance (fire, mining) on ant and small mammal species diversity in Australian heathland. *Proceedings of*

the Symposium on Dynamics and Management of Mediterranean-Type Ecosystems, June 22-26, 1981, San Diego, California, San Diego State University. United States Department of Agriculture Forest Service Pacific Southwest Forest and Range Experiment Station General Technical Report PSW-58, 213-219.

- Gallego, I., Davidson, T.A., Jeppesen, E., Pérez-Martínez, C., Sánchez-Castillo, P., Juan, M., Fuentes-Rodríguez, F., León, D., Peñalver, P., Toja, J. & Casas, J.J. (2012). Taxonomic or ecological approaches? Searching for phytoplankton surrogates in the determination of richness and assemblage composition in ponds. *Ecological Indicators*, 18, 575-585.
- Gibb, H., Sanders, N.J., Dunn, R.R., Arnan, X., Vasconcelos, H.L., Donosco, D.A., Andersen, A.N., Silva, R.R., Bishop, T,R., Gomez, C., Grossman, B.F., Yusah, K.M., Luke, S.H., Pacheco, R., Pearce-Duvet, J., Retana, J., Tista, M. & Parr, C.L. (2018). Habitat disturbance selects against both small and large species across varying climates. *Ecography*, 41, 1184-1193.
- Gómez, C., Casellas, D., Oliveras, J. & Bas, J.M. (2003). Structure of ground-foraging ant assemblages in relation to land-use change in the northwestern Mediterranean region. *Biodiversity and Conservation*, 12, 2135-2146.
- Grandin, U., Lenoir, L. & Glimskär, A. (2013). Are restricted species checklists or ant communities useful for assessing plant community composition and biodiversity in grazed pastures? *Biodiversity Conservation*, 22, 1415-1434.
- Grantham, H.S., Pressey, R.L., Wells, J.A. & Beattie, A.J. (2010). Effectiveness of biodiversity surrogates for conservation planning: different measures of effectiveness generate a kaleidoscope of variation. *PloS ONE*, 5(7), e11430.
- Greenslade, P.J.M. (1964). Pitfall trapping as a method for studying populations of Carabidae (Coleoptera). *Journal of Animal Ecology*, 33(2), 301-310.

- Greenslade, P.J.M. (1973). Sampling ants with pitfall traps: digging-in effects. *Insectes Sociaux*, 20(4), 343-353.
- Greenslade, P.J.M. (1976). The meat ant Iridomyrmex purpureus (Hymenoptera: Formicidae) as a dominant member of ant communities. Journal of the Australian Entomology Society, 15, 237-240.
- Heink, U. & Kowarik, I. (2010). What are indicators? On the definition of indicators in ecology and environmental planning. *Ecological Indicators*, 10, 584-598.
- Hoffmann, B.D. & Andersen, A.N. (2004). *Ant Monitoring at Callide Mine, Central Queensland. 2004 Sampling*. A consultancy report prepared for Callide Mine. CSIRO Sustainable Ecosystems, Darwin.
- IUCN (2001). *IUCN Red List Categories and Criteria version 3.1*. IUCN Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, UK.
- Jackson, G.P. & Fox, B.J. (1996). Comparison of regeneration following burning, clearing or mineral sand mining at Timago, NSW: II. Succession of ant assemblages in a coastal forest. *Australian Journal of Ecology*, 21, 200-216.
- Kaspari, M. & Majer, J.D. (2000). Using ants to monitor environmental change. In Agosti, D., Majer, J.D., Alonso, L.E. & Schultz, T.R. (Eds.) *ANTS: Standard methods for measuring and monitoring biodiversity*, Smithsonian Institute Press, Washington, 89-98.
- Kaspari, M., O'Donnel, S. & Kercher, J.R., (2000). Energy, density and constraints to species richness: ant assemblages along a productivity gradient. The American Naturalist, 155(2), 280-293.
- Kindsvater, H.K., Dulvy, N.K., Horswill, C., Juan-Jordá, MJ., Mangel, M. & Matthiopoulos, J. (2018). *Trends in Ecology and Evolution*, 33(9), 676-688.
- Koo, T.K. & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155-163.

- Kwon, T.S. & Lee, C.M. (2015). Prediction of abundance of ants according to climate change scenarios RCP 4.5 and 8.5 in South Korea. Journal of Asia-Pacific Biodiversity, 8, 49-65.
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159-174.
- Lattke, J.E. (2000). Specimen processing: building and curating an ant collection. In Agosti, D., Majer, J.D., Alonso, L.E. & Schultz, T.R. (Eds.) *ANTS: Standard methods for measuring and monitoring biodiversity*, Smithsonian Institute Press, Washington, 155-171.
- Lawes, M.J., Moore, A.M., Andersen, A.N., Preece, N.D., Franklin, D.C. (2017). Ants as ecological indicators of rainforest restoration: community convergence and the development of an ant forest indicator in the Australian wet tropics. *Ecology and Evolution*, 7, 8442-8455.
- Lee, K.M., Lee, J., Chung, C.Y., Ahn, S., Sung, K.H., Kim, T.W., Lee, H.J. & Park, M.S. (2012). Pitfalls and important issues in testing reliability using intraclass correlation coefficients in Orthopaedic research. Clinics in Orthopedic Surgery, 4, 149-155.
- Longino, J.T. (2000) What to do with the data. In Agosti, D., Majer, J.D., Alonso, L.E. & Schultz, T.R. (Eds.) *ANTS: Standard methods for measuring and monitoring biodiversity*, Smithsonian Institute Press, Washington, 186-203.
- Lovell, S., Hamer, M., Slotow, R. & Herbert, D. (2007). Assessment of congruency across invertebrate taxa and taxonomic levels to identify potential surrogates. *Biological Conservation*, 139, 113-125.
- Majer, J.D. (1983) Ants: bio-indicators of minesite rehabilitation, land-use and land conservation. *Environmental Management*, 7(4), 375-383.
- Majer, J.D. (1985). Seasonality of ants in south-western Australia. WAIT School of Biology Bulletin, 12, 23-26.
- Majer, J. D. (1989). Long-term colonization of fauna in reclaimed lands. In Majer, J. D. (Ed.) Animals in primary succession: The role of fauna in reclaimed lands, Cambridge University Press, Cambridge, 143-174.

- Majer, J. D. (1992). Ant recolonization of rehabilitated bauxite mines of Pocos de Caldas, Brazil. Journal of Tropical Ecology, 8, 97-108.
- Majer, J.D. (1996). Ant recolonization of rehabilitated bauxite mines at Trombetas, Pará, Brazil. Journal of Tropical Ecology, 12, 257-273.
- Majer, J.D. (1997). The use of pitfall traps for sampling ants – A critique. *Memoirs of Museum Victoria*, 56(2), 323-329.
- Majer, J.D., Sartori, M., Stone, R., Perriman, W.S. (1982). Recolonisation by ants and other invertebrates in rehabilitated mineral sand mines near Eneabba, Western Australia. *Reclamation and Revegetation Research*, 1, 63-81.
- Majer, J.D., Nichols, O.G. (1998). Long-term recolonization patterns of ants in Western Australian rehabilitated bauxite mines with reference to their use as indicators of restoration success. *Journal of Applied Ecology*, 35(1), 161-182.
- Majer, J.D., Brennan, K.E.C., Moir, M.L. (2007). Invertebrates and the restoration of a forest ecosystem: 30 years of research following Bauxite mining in Western Australia. *Restoration Ecology*, 15(4(Supplement)), S104-S115.
- Majer, J.D., Heterick, B., Gohr, T., Highes, E., Mounsher & L., Grigg, A. (2013). Is thirty-seven years sufficient for full return of the ant biota following restoration? *Ecological Processes*, 2(1), 19-30.
- Marsh, A.C. (1984). The efficacy of pitfall traps for determining the structure of a desert ant community. *Journal of the Entomological Society of Southern Africa*, 47(1), 115-120.
- Marshall, J.C., Steward, A.L. & Harch, B.D. (2006). Taxonomic resolution and quantification of freshwater macroinvertebrate samples from an Australian dryland river: the benefits and costs of using species abundance data. *Hydrobiologia*, 572, 171-194.
- McGraw, K.O. & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30-46.

- McGraw, K.O. & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1(1), 30-46.
- McHugh, M.L. (2012) Interrater reliability: the kappa statistic. Biochemica Medica, 22(3), 276-282.
- Minerals Council of Australia (2017). *Submission to Senate Environment and Communications Reference Committee Enquiry into the rehabilitation of mining and resources projects as it relates to Commonwealth responsibilities.* (accessed 04/09/2018 from https://www.minerals.org.au/news/mca_submission_to_the_senate_inquiry_into_mine_rehabilitation).
- Nakamura, A., Proctor, H. & Catterall, C.P. (2003) Using soil and litter arthropods to assess the state of rainforest restoration. *Ecological Management and Restoration*, 4(Supplement), S20-S28.
- Nakamura, A., Catterall, C.P., House, A.P.N., Kitchung, R.L. & Burwell, C.J. (2007) The use of ants and other soil and litter arthropods as bio-indicators of the impact of rainforest clearing and subsequent land use. *Journal of Insect Conservation*, 11, 177-186.
- Ness, J.H., Bronstein, J.L. Andersen, A.N. & Holland, J.N. (2004). Ant body size predicts dispersal distance of ant-adapted seeds: implications of small-ant invasions. *Ecology*, 85(5), 1244-1250.
- Nichols, O.G. & Nichols, F.M. (2003). Long-term trends in faunal recolonisation after bauxite mining in the Jarrah Forest of Southwestern Australia. *Restoration Ecology*, 11(3), 261-272.
- OECD – Organisation for Economic Co-Operation Development (2003). *OECD Environmental Indicators. Development, Measure and Use*. OECD, Paris.
- Olson, D.M. (1991). A comparison of the efficacy of litter sifting and pitfall traps for sampling leaf litter ants (Hymenoptera, Formicidae) in a tropical wet forest, Costa Rica. *Biotropica*, 23(2), 166-172.
- Ottonetti, L., Tucci, L., Santini, G. (2006). Recolonization patterns of ants in a rehabilitated lignite mine in central Italy: potential for the use of

Mediterranean ants as indicators of restoration processes. *Restoration Ecology*, 14(1), 60-66.

- Pandley B., Agrawal, M. & Singh, S. (2014). Effects of coal mining activities on soil properties with special reference to heavy metals. In Raju, N.J. (Ed.) *Geostatistical and Geospatial Approaches for the Characterization of Natural Resources in the Environment: Challenges, Processes and Strategies* (2016). Publishing Company, New Delhi, 369-372.
- Pik, A.J., Oliver, I. & Beattie, A.J. (1999). Taxonomic sufficiency in ecological studies of terrestrial invertebrates. *Australian Journal of Ecology*, 24, 555-562.
- Postle, A.C. & Philips, R. (n.d.) A survey of ground-dwelling invertebrate fauna on the Argyle diamond lease. School of Environmental Biology, University of Technology, Perth, Western Australia.
- Ribas, C.R., Schmidt, F.A., Solar, R.R.C., Campos, R.B.F., Valentim, C.L. & Schoereder, J.H. (2012). Ants as indicators of the success of rehabilitation efforts in deposits of gold mining tailings. *Restoration Ecology*, 20(6), 712-720.
- Ribas, L.G.d.S. & Padial, A.A. (2015). The use of coarser data is an effective strategy for biological assessments. *Hydrobiologia*, 747, 83-95.
- Roberto, C. & Brandão (2000). Major Regional and type collections of ants (Formicidae) of the world and sources for the identification of ant species. In Agosti, D., Majer, J.D., Alonso, L.E. & Schultz, T.R. (Eds.) *ANTS: Standard methods for measuring and monitoring biodiversity*, Smithsonian Institute Press, Washington, 172-185.
- Rosser, N. & Eggleton, P. (2012). Can higher taxa be used as a surrogate for species-level data in biodiversity surveys of litter/soil insects? *Journal of Insect Conservation*, 16, 87-92.
- Siddig, A.A.H., Ellison, A.M., Ochs, A., Villar-Leeman, C., Lau, M.K. (2016). How do ecologists select and use indicator species to monitor ecological change? Insights from 14 years of publication in *Ecological Indicators*. *Ecological Indicators*, 60, 223-230.

- Souza, J.L.P., Baccaro, F.B, Pequeno, P.A.C.L., Franklin, E. & Magnusson, W.E. (2018). Effectiveness of genera as a higher-taxon substitute for species in ant biodiversity analyses is not affected by sampling technique. *Biodiversity and Conservation*, DOI: 10.1007/s10531-018-1607-x
- Stacey, K. (2003). *Callide Mine Ant Monitoring Program 2003.* Callide Mine Environment Department, Callide.
- Thorne, R.St.J., Williams, W.P. & Cao, Y. (1999). The influence of data transformations on biological monitoring studies using macroinvertebrates. *Water Research*, 33(2), 343-350.
- van Hamburg, H., Andersen, A.N., Meyer, W.J. & Robertson, H.G. (2004). Ant community development on rehabilitated ash dams in the South African Highveld. *Restoration Ecology*, 12(4), 552-558.
- Vellend, M., Lilley, P.L. & Starzomski, B.M. (2008). Using subsets of species in biodiversity surveys. *Journal of Applied Ecology*, 45, 161-169.
- Whitford, W.G. & Ettershank, G. (1975). Factors affecting foraging activity in Chichuahuan desert harvester ants. Environmental Entomology, 4, 689-696.
- Wittkuhn, R.S., McCaw, L., Willis, A.J., Robinson, R., Andersen, A.N., Van Heurck, P., Farr, J., Liddelow, G. & Cranfield, R. (2011). Variation in fire interval sequence has minimal effects on species richness and composition in fire-prone landscapes of south-west Western Australia. *Forest Ecology and Management*, 261, 965-978.
- Woinarski, J.C.Z., Andersen, A.N., Churchill, T.B. & Ash, A.J. (2002). Response of ant and terrestrial spider assemblages to pastoral and military land use, and to landscape position, in a tropical savanna woodland in northern Australia. *Austral Ecology*, 27, 324-333.
- Wright, I.A., Chessman, B.C., Fairweather, P.G. & Benson, L.J. (1995). Measuring the impact of sewage effluent on the macroinvertebrate community of an upland stream: the effect of different levels of taxonomic resolution and quantification. *Australian Journal of Ecology*, 20, 142-149.